RELIABILITY AND DIAGNOSTIC VALIDITY OF CLINICAL EXAMINATION AND PATIENT SELF-REPORT MEASURES IN CARPAL TUNNEL SYNDROME

A Thesis Submitted to the College of

Graduate Studies and Research

in Partial Fulfillment of the Requirements

for the Degree Master of Science

in the Department Community Health and Epidemiology

University of Saskatchewan

Saskatoon

By

Brenna Bath

Permission to Use

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Community Health and Epidemiology
University of Saskatchewan
Saskatoon, Saskatchewan S7N 5E5

Abstract

Study Design: A blinded, prospective diagnostic test study was conducted.

Objectives: To assess the inter-tester reliability of clinical examination items for carpal tunnel syndrome (CTS), to assess the validity of individual clinical test items and self-report measures for the diagnosis of CTS and to assess the accuracy of an optimum test item cluster for the diagnosis of CTS.

Study Rational: Examination of the diagnostic validity of various clinical tests for the diagnosis of CTS has shown mixed results and the reliability of many of these tests has not been determined. The majority of the diagnostic validity research for CTS has examined tests individually which is in contrast to clinical practice where the results of multiple tests are combined as part of the clinical reasoning process in order to formulate a differential diagnosis: the test item cluster, derived through logistic regression, is proposed as a means to integrate the validity of multiple tests.

Methods: 37 subjects (74 "hands") were recruited from a convenience sample of consecutive patients referred to for electrodiagnostic (EDX) testing due to upper extremity symptomatic complaints. Subjects underwent EDX testing followed by completion of self-report questionnaires and a standardized clinical examination by examiners blinded to EDX results. Diagnostic validity was determined for both "general" and "restricted" CTS classification groupings.

Results: Out of 18 clinical test items, 12 had reliability coefficients (i.e. ICC or Kappa) of .40 or greater. There were 10 clinical exam and self-report items that were found to have likelihood point estimates above 2 or below 0.50 for the general diagnostic classification and 6 items had acceptable validity for the restricted classification. The test item cluster (TIC) derived for the general classification included "hand numbness" and symptom reproduction with the upper limb neurodynamic test 1. The TIC derived for the restricted classification included "hand numbness" and the overall score of the symptom component of the Bringham Carpal Tunnel Questionnaire. The 95% confidence intervals for most likelihood ratio point estimates were wide.

Conclusions: The TICs for both classification groupings did not yield improved diagnostic validity beyond that found with the single best test item ("hand numbness"). The value of the single best test item "hand numbness" was in a negative response.

Further investigation is required to validate the TIC and the single best test item and to improve the point estimate precision.

Acknowledgements

My thanks and gratitude to my thesis supervisors, Dr. L. Tan and Dr. B. Janzen: both allowed me a great deal of independence and yet provided me with excellent and timely feedback and advice when I needed it.

My thanks as well to the remainder of my thesis committee: Dr. N. Muhajarine, Dr. L. Harrison, Dr. R. Li Pi Shan and Dr. N. Ashworth. Dr. Muhajarine joined my committee mid way through the process and did a great job as committee chair. Dr. Harrison provided insightful comments throughout the process and has been a wonderful source of encouragement. Dr. Li Pi Shan served a dual role as the Physiatrist involved in the data collection and as a committee member; his relaxed and easygoing manner in both roles was well appreciated. Dr. Ashworth's thoroughness and interest in the topic proved him to be the ideal external reviewer for this thesis.

Thanks to John Marshall who devoted a great deal of time to assisting me develop the operational definitions for the clinical tests and who served as the second tester during the data collection periods.

Thank you to the University of Saskatchewan's School of Physical Therapy for funding support and to the Saskatoon Health Region for provision of personnel and space required for completion of this study.

Finally, thank you to my husband Jon for his ongoing tech support.

Dedication

This thesis is dedicated to my husband Jon and my son William. You have both been an ongoing source of joy and serve as constant reminders to me that there is life outside of research and work.

Table of Contents

List of Abbreviations

APB- Abductor pollicus brevis

BCTQ- Bringham Carpal Tunnel Questionnaire

BCTQ-FN- Bringham Carpal Tunnel Questionnaire Function Section

BCTQ-SSS- Bringham Carpal Tunnel Questionnaire Symptom Severity Scale

BMI- Body Mass Index

CI- Confidence interval

CTS- Carpal Tunnel Syndrome

DASH- Disability of Arm Shoulder and Hand Questionnaire

DOR- Diagnostic odds ratio

EDX- Electrodiagnostic test

G- General

HL- Hosmer Lemeshow

ICC- Intraclass correlation coefficient

LR- Likelihood Ratio

NSAID- Non-steroidal anti-inflammatory drug

PT- Physical Therapist

R- Restricted

ROC curve- Receiver operator characteristic curve

ROM- Range of motion

SEM- Standard error of measurement

SGI- Saskatchewan Government Insurance

SSS- Symptom severity scale

TIC- Test item cluster

ULNT- Upper limb neurodynamic test

ULNT 1- Upper limb neurodynamic test with a median nerve bias

ULNT 1 >10 deg diff- Upper limb neurodynamic test with a median nerve bias, greater than 10 degrees difference elbow extension between right and left upper extremities

ULNT 1 SF provocation- Upper limb neurodynamic test with a median nerve bias, symptom provocation or relief with cervical side flexion

WCB- Worker's Compensation Board

List of Tables

List of Figures

1. INTRODUCTION

1.1 Purpose of Thesis Study

- To assess the inter-tester reliability of clinical examination items for the diagnosis of carpal tunnel syndrome.
- To assess the validity of individual clinical examination items and self-report items for the diagnosis of carpal tunnel syndrome.
- To assess the validity of an optimum test item cluster for the diagnosis of carpal tunnel syndrome.

1.2 Literature Review

1.2.1 Carpal Tunnel Syndrome

1.2.1.1 *CTS Description:*

Carpal tunnel syndrome (CTS) can be a cause of pain and functional impairment of the hand due to compression of the median nerve at the wrist.[1,2] Usual symptoms include numbness, tingling and pain predominantly in the median nerve distribution of the hand; however, the symptoms can frequently be present in all fingers of the hand[3] or proximally in the forearm. The symptoms may or may not be accompanied by objective changes in sensation and strength of median-innervated structures in the hand. [4]

1.2.1.2 *CTS Frequency:*

The reported frequency of CTS is variable and is dependent on the type of criteria used to identify the presence of the condition and the population studied. The prevalence has been estimated to 50 - 160 cases per 1000 subjects in the general population [5,6,7] and up to 500 cases per 1000 subjects in certain high risk groups.[8] Incidence may be as anywhere from 2.5 cases per 1000 subjects[9] to as high as 150 per 1000 subjects per year.[10]

1.2.1.3 *CTS Associated Factors:*

Carpal tunnel syndrome is associated with numerous epidemiological risk factors including genetic, medical, social, vocational, avocational and demographic. However, definitive causal factors for CTS have yet to be identified. Women are much more susceptible than men to develop CTS; the female to male ratio has been reported to be from 3 to 10 : 1. [6,11] The peak age of development for CTS is 45-60 years [6] with an estimated 10% of people with CTS younger than 31 years. [6,9] Caucasians are more likely than African Americans to develop CTS. [10,12] Increased body mass index (BMI) (especially a recent increase), short stature, dominant hand and square shaped wrist are other factors associated with CTS.[9] A strong familial susceptibility for CTS, related to inherited genetic characteristics and medical conditions, may exist. [9] Medical conditions that may be related to the development of CTS include history of a colles fracture (fracture of distal radius), space occupying lesions in the carpal tunnel (e.g. flexor tenosynovitis, ganglions, hemorrhage, edema), diabetes, rheumatoid arthritis, osteoarthritis of the carpo-metacarpal joint of the thumb and recent menopause.[9,11] Vocational and avocational activities with prolonged severe force through the wrist, high amounts of repetitive movements, prolonged extreme factors and exposure to vibration and/or cold may also be associated with the development of CTS. [9,11] Other associated risk factors include: pregnancy, use of a walking aid or wheelchair, smoking and alcohol use. [11]

1.2.1.4 *CTS Treatment:*

Numerous treatment options exist for the management of CTS symptoms. Mild symptoms of short duration are thought to be self-limiting and will often resolve with conservative measures (i.e. splinting, non-steroidal anti-inflammatory drugs (NSAIDs), activity modification). Clinical management of more severe cases may include local corticosteroid injection or surgical release of the transverse ligament.

Clinical trials examining the effectiveness of various management options show variable results. Surgical treatment has been shown to be more effective than splinting in moderate to severe cases of CTS. [13] Surgical release of the median nerve, however, can have a 15-20% failure rate with variable long term success rates [14] and the potential for

complications. [15,16] Corticosteroid injection is demonstrated to give improved symptom relief over placebo or NSAIDs in the short term; however, it shows no difference in benefit compared to splinting and NSAID's beyond 8 weeks.[17] Conservative measures such as splinting, ultrasound, yoga and carpal bone mobilization[18] have shown significant short term improvement of CTS symptoms; however long term effectiveness has yet to be demonstrated and the effectiveness of these measures compared to surgical intervention has yet to be determined. Neurodynamic mobilization is an emerging conservative treatment technique that shows some promise for symptomatic management of carpal tunnel syndrome.[19,20] Given the variable presentation and potentially complicated nature of CTS, ultimately a multimodal approach to treatment will likely be most effective; this type of approach has yet to be formally examined.

1.2.1.5 *CTS Other Considerations:*

The incidence of bilateral CTS has been reported to be up to 87% and is often associated with symptoms of longer duration.[21] There is also evidence of frequent subclinical median nerve compression in contralateral limbs of patients with unilateral CTS.[22] The high frequency of bilateral symptoms suggests that median nerve compression distally at the wrist is not likely the sole contributor to CTS symptoms; the double crush hypothesis may partially explain this phenomenon.

The double crush hypothesis states that a proximal source of nerve compression or stretch would make the distal nerve more susceptible to compression.[23] This clinically based hypothesis came from observations of coexistence of cervical discogenic disease with CTS. Recent research using MRI has demonstrated a higher incidence of narrowed cervical foraminae in CTS patients and a concordance with affected nerve roots on the same side as the CTS, which supports the double crush hypothesis.[24]

1.2.2 **Neurodynamics**

1.2.2.1 *Definition:*

Neurodynamics refers to the mechanics and physiology of the nervous system and how they relate to each other. The nervous system exists in physiological and

mechanical continuity.[25] For example, mechanical stresses applied to nerves evoke physiological responses such as alterations in axonal transport or intraneural blood flow; conversely, any physiological disturbance of the nerve, such as with diabetes, predisposes it to mechanical disturbances.[26]

1.2.2.2 *Neurodynamics' relationship to CTS:*

Restricted longitudinal movement of the median nerve in patients with CTS has been documented via medical imaging. [27, 28] Limited adaptation of nerve structures to changes in length of the nerve path during upper limb movement could eventually be a source of neurological symptoms and lead to secondary alterations in nerve function.[29] Physiological sliding could be impaired when the median nerve is trapped at any point along its course. Impaired longitudinal sliding should be considered among the mechanisms contributing to electrophysiological abnormalities in patients with CTS. [29]

1.2.2.3 *Upper Limb Neurodynamic Tests (ULNT):*

A neurodynamic test is a sequence of movements designed to assess the mechanics and physiology of part of the nervous system by elongating the length of the nerve and associated connective tissue and by increasing the pressure in and around the peripheral nerve.[30] The ULNT was initially conceptualized by Elvey[31] as a physical examination of the neural tissues of the upper quadrant in the investigation of arm pain and regional pain syndromes of the upper quarter. The ULNT provides a simple, effective and repeatable method of clinical examination by moving the neural tissues and stimulating them mechanically in order to gain an impression of their mobility and sensitivity to mechanical stresses.[32] The test may elicit a component of neuromechanical tension, increased mechanosensitivity or a disturbance of pain mechanisms. Neurodynamic testing with a median nerve bias (ULNT1) has been proposed as a potentially valid indicator of pathology along the pathway of the median nerve [30] but the clinical accuracy of this test for use in CTS has yet to be determined.

1.2.3 **Carpal Tunnel Syndrome Diagnosis**

1.2.3.1 *Electrodiagnostic (EDX) Testing*

The goal of electrodiagnostic testing for CTS is to determine whether median nerve function is impaired at the wrist. EDX studies (nerve conduction studies and electromyography) can include an examination of both sensory and motor nerves. There is currently no widely agreed upon consensus regarding which EDX study technique or combination of techniques is best for detecting CTS [33]; however, median nerve sensory conduction measures appear to be preferred (if the nerve is intact). [5]

EDX studies are often considered to be the most accurate and one of the few "objective" tests to document the presence of CTS. Other, less common, diagnostic tests include ultrasound and magnetic resonance imaging (MRI). EDX studies are far from a perfect "gold standard" for CTS diagnosis as false positives and negatives findings with this method of testing have been well documented. [33] False negatives can occur because of the intermittent nature of the syndrome or because symptoms emanate from small unmyelinated fibres that are invisible to surface electrodes (EDX only detects larger myelinated fibers). Carpal tunnel release surgery can be performed successfully in patients with normal EDX findings. In patients with positive EDX findings who undergo surgery, their symptoms usually resolve within days despite EDX abnormalities that persist for longer. [1]

1.2.3.2 *CTS Symptom Presentation*

CTS is, by definition, a clinical syndrome with a characteristic symptom presentation. The main symptoms of CTS are paresthesia (numbness and/or tingling) and pain of the fingers and hand. Usually the symptoms are in the median nerve distribution (Figure 1), but frequently complaints involve all fingers and/or symptoms proximal to the wrist.[3]

Figure 1: Median Nerve Distribution

The presence of "typical" median nerve distributed symptoms does often not necessarily correlate with EDX evidence of median nerve pathology.[34] Reliance on symptom presentation alone for the diagnosis of CTS is, therefore, not ideal.

1.2.3.3 *Recommended Diagnostic Criteria*

The combination of electrodiagnostic study findings and symptom characteristics likely provides the most accurate CTS diagnosis according to a consensus group of medical researchers with experience conducting epidemiological studies of CTS. [33] This group combined the symptom presentation gathered from the use of hand diagrams to label "classic/probable", "possible" and "unlikely" symptom characteristics (see Table 1) with EDX study findings.

Table 1: Classification of Symptom Quality and Location (From Rempel, 1998)[33]

Symptom	Description
Classic/probable	Numbness, tingling, burning or pain in at least 2 of digits 1, 2 or 3. Palm pain, wrist pain, or radiation proximal to the wrist is allowed
Possible	Tingling, numbness, burning or pain in at least 1 of digits 1, 2 or 3
Unlikely	No symptoms in digits 1, 2, and 3

They were unable to come to a consensus on the likelihood of CTS with a combination of "classic/probable" symptom presentation and negative EDX studies (see Table 2). They assigned an "ordinal likelihood of CTS" based on the relative diagnostic strength of the symptom presentation combined with the EDX results. The group determined that the addition of physical examination findings add little diagnostic value if EDX findings and

Table 2: Estimation likelihood of CTS for Case Definitions of CTS that include EDX studies (Adapted from Rempel, 1998)[33]

Symptom	EDX	Ordinal likelihood of CTS	Interpretation**
Classic/probable	Positive	+++	i.e. High likelihood of CTS
Possible	Positive	++	i.e. High likelihood of CTS
Classic/probable	Negative	+/-*	i.e. Mixed likelihood of CTS
Possible	Negative	-	i.e. Low likelihood of CTS
Unlikely	Positive	-	i.e. Low likelihood of CTS
Unlikely	Negative	--	i.e. Low likelihood of CTS

* ***No consensus achieved on whether likelihood should be – or +.***
***** Thesis author's interpretation.***

symptom characteristics are available; however, in the absence of EDX findings, combinations of physical examination findings and symptom characteristics provide the greatest diagnostic information.

1.2.4 **Physical Examination**

1.2.4.1 Types of Validity

Validity is the extent to which a meaningful interpretation can be inferred from a measurement. [35] There are several different types of validity including construct, content, and criterion validity. Constructs are artificial frameworks which are not directly observable.[36] Construct validity is often determined through logical argument rather than

experimental study. Construct validation is the theoretical foundation on which all other types of validity depend. [30] Content validity deals with the concept of how well the measure reflects or samples the construct. The criterion validation of a measure is determined by comparing it with an accepted standard of measurement. Diagnostic validity can be considered a type of criterion validity.

1.2.4.2 *Problems with Determining Diagnostic Validity*

A plethora of research has been undertaken to examine the diagnostic validity of various physical examination and clinical tests for CTS; fewer studies have examined the reliability of these tests (See Table 3). The results of these studies are extremely variable and, because of the use of different criteria standards and control groups, comparison among studies is difficult.

Studies that use asymptomatic patients as controls rather than patients with other upper extremity pathologies create an artificial simplicity in diagnosis that does not challenge the diagnostic test nor reflect clinical practice, as the role of the clinical test is to distinguish CTS patients from those with other pathologies that might be potentially confused with CTS. [37] Sackett [38] reports that inclusion of asymptomatic controls can overestimate the true accuracy of a test used in clinical evaluation.

Comparability among studies is also difficult because of the use of differing "gold standards." A main limitation in the literature is the lack of an ideal criteria standard which complicates all research in CTS. One study may use the results of EDX testing only [39], another may use a combination of EDX findings and symptoms [40], and another yet may use success with surgery as evidence of CTS.[41]

Table 3: Validity and Reliability of Physical Examination Procedures for CTS*

Physical Examination Test	Validity	Reliability
Phalen's[42]	Sensitivity 67-88% Specificity 32-86%	Intratester kappa=0.53 Intertester kappa=0.65
Tinel's[42]	Sensitivity26-73% Specificity 55-94%	Intratester kappa=0.80 Intertester kappa=0.79
Carpal compression test[42]	Sensitivity 87% Specificity 90-95%	n/a
Semmes-Weinstein pressure sensibility[42]	Sensitivity26-73% Specificity 55-94%	Intratester ICC**=0.15 Intertester ICC =0.71
Sharp/Dull sensibility[42]	Sensitivity 15-93% Specificity 51-85%	n/a
Thumb abduction strength (APB)***[40]	Sensitivity 66% Specificity 66%	n/a
Grip strength	Sensitivity 48% Specificity 30-38% [41]	Reliability coefficient=0.88 [42]
Key pinch strength	Sensitivity 33% Specificity 52-61% [41]	Intratester Pearson r=0.98 Intertester Pearson r=0.98 [42]
Tip pinch strength	Sensitivity 65% Specificity 38-41% [41]	Intratester Pearson r=0.98 Intertester Pearson r=0.98 [40]
Tripod pinch strength	Sensitivity 43% Specificity 49% [41]	Intratester Pearson r=0.98 Intertester Pearson r=0.98 [42]
ULNT 1	n/a	Intratester ICC=0.98 [43]
Flick Sign [1]	Sensitivity 93% Specificity 96%	n/a
Thenar eminence wasting[42]	Sensitivity 3-19% Specificity 93-100%	n/a
Square shape wrist [40]	Sensitivity 69% Specificity 73%	n/a

**See Appendix A for description of each clinical test and for operational definitions.*

*** ICC = intraclass correlation coefficient*

****APB = abductor pollicis brevis*

1.2.5 **Self-Report Measures**

Self-report measures are emerging as important tools for outcome measurement in both clinical environments and with intervention studies such as controlled trials.

1.2.5.1 *Disability of Arm Shoulder and Hand Questionnaire*

The DASH (Disability of Arm Shoulder and Hand questionnaire- Appendix A) is an outcome assessment tool of self-perceived disability.[44] The DASH has been proposed as a valid tool for CTS [45]; however, its diagnostic validity for CTS has not been examined.

1.2.5.2 *Bringham Carpal Tunnel Questionnaire*

A self-administered questionnaire (Bringham Carpal Tunnel Questionnaire- BCTQ- Appendix B) for the assessment of severity of symptoms and functional status in patients who have CTS was developed by Levine et al.[46] The measurement scales have been shown to be highly reproducible, internally consistent and responsive to clinical change. [46] The diagnostic validity of this tool for CTS has yet to be extensively examined.

1.3 **Interpretation of Diagnostic Tests: Determining Validity**

There are various methods for the interpretation of diagnostic tests: sensitivity, specificity, predictive values, likelihood ratios and diagnostic odds ratios (refer to Appendix C for brief descriptors of each and formulae for calculating). The use of sensitivity and specificity is far more common than the other measures; recently, however, likelihood ratios are being recommended as more clinically relevant. [47, 48]

Predictive values appear on the surface to be a useful and simple means of interpreting clinical test validity: "given a test result (positive or negative) what is the probability that the result is correct?".[47] Predictive values can, however, be deceptive because they are highly dependent on the prevalence of the condition of interest in the sample. Positive predictive values will be lower and negative predictive values will be higher in samples with a low prevalence of the condition. The trends reverse if the prevalence is high.[49]

Prevalence levels do not affect sensitivity and specificity in this manner.[38] Sensitivity and specificity values are useful for interpreting clinical tests. Sackett et al [38] advocate the use of the following acronyms to assist with ease of interpretation: "SnNout" (if sensitivity (Sn) is high, a negative (N) result is useful in ruling out (out) the condition) and "SpPin" (if specificity (Sp) is high, a positive (P) result is useful in ruling in (in) the condition).

Even though sensitivity and specificity values provide useful information, several shortcomings have been identified.[47] These values tend to work in the opposite direction of clinical decision making: clinicians have knowledge of the test result and want to infer the probability that a test result is correct, where as sensitivity and specificity values infer the probability of a correct test given the result of the reference standard. Sensitivity and specificity values can only be used independently as estimates of the accuracy of negative or positive test results, but this information cannot be combined for interpretation. The performance of a diagnostic test, in the context of predictive values, is dependent on the pre-test probability that the condition is present.[48] Useful tests should produce large shifts in probability (i.e. a large change in pre-test probability to post-test probability) once the result of the test is known.[50,51] Sensitivity and specificity values cannot be used to quantify shifts in probability of the condition.[47]

The use of likelihood ratios (LRs) overcomes the aforementioned shortcomings. LRs combine the sensitivity and specificity values into a ratio that can be used to quantify shifts in probability once the diagnostic test results are known.[52] The positive LR indicates the increase in odds favoring the condition when the test is positive. Tests with large positive LRs generally have a high specificity because both values reflect the usefulness of a positive test. The negative LR indicates the decrease in odds favoring the condition given a negative test result; a small negative LR means that a test is useful for ruling out a condition when negative (small negative LRs usually correspond to high sensitivity values). Refer to Table 4 for a guide to LR interpretation.

Table 4: Interpretation of LR values[47]

Positive LR	Negative LR	Interpretation
>10	<0.1	Generate large and often conclusive shifts in probability
5-10	0.1-0.2	Generate moderate shifts on probability
2-5	0.2-0.5	Generate small but sometimes important shifts in probability
1-2	0.5-1	Alter probability to a small, and rarely important, degree

Likelihood ratios can be used as part of the clinical reasoning process to determine the post-test probability of a condition given a known or estimated pre-test probability (i.e. prevalence). This can be estimated from a nomogram or calculated via a mathematical formula (Appendix C).[53]

The diagnostic odds ratio (DOR) is a relatively new, and somewhat controversial, means of interpreting diagnostic tests.[54] The use of paired indicators, such as positive and negative LRs, in comparing the performance of competing tests can be a disadvantage if one test does not outperform the other on both indicators.[55] The DOR combines the positive and negative LR values of a test into one performance indicator which facilitates test comparison (see Appendix C for DOR formula). A disadvantage of the DOR is that, as a combined measure, it does not give an indication of the relative value or contribution of a positive or negative test result, thus the clinical usefulness is limited.

1.4 **Rationale and Significance of Study:**

EDX studies have been touted as the "gold standard" for clinical evidence of CTS. Reliance on EDX alone for diagnosis, even when combined with symptom presentation, however, does have several drawbacks. EDX tests require expensive equipment and specialized personnel. They can also be relatively time consuming and uncomfortable. Clinical tests and self-report questionnaires offer the advantage of being simple, inexpensive and quick to perform. Clinical testing and questionnaires can potentially be used as a quick screen to determine the likelihood of CTS in the presence

of symptoms and contributing historical factors. If a positive likelihood of CTS is determined with clinical testing and questionnaires appropriate interventions could be undertaken prior to EDX testing. Appropriate interventions could, therefore, be potentially initiated earlier; particularly in cases where treatment is delayed until confirmation of a diagnosis with EDX testing is obtained.

Few studies have examined the reliability of commonly used physical examination procedures for CTS diagnosis. The reliability of a test is an important precursor to determining its validity. For example, if a test does not produce consistent and repeatable results, then the degree to which a meaningful interpretation can be inferred from a test is limited.

Research to date on clinical test validity for CTS has primarily been done on localized tests of provocation at the wrist (e.g. Phalen's, Tinel's) or impairment measures (e.g. sensation, strength). The diagnostic validity of the ULNT1 and the BCTQ has been examined by a recently published study [56]; however the DASH has yet to be examined in this regard.

The majority of the validity research for CTS has examined clinical tests individually. This is in contrast with clinical practice where the results of multiple tests are combined as part of the clinical reasoning process in order to formulate a differential diagnosis. This study proposes to examine the results of multiple tests individually and in combination in order to determine if an optimal test cluster can be derived. The development of clinical prediction rules[57], or test item clusters (TICs), is a recent trend in the literature. Determination of a valid test item cluster (TIC) is a means of combining multiple tests in a clinically meaningful way. A valid TIC would more accurately reflect common clinical practice and, hopefully, provide clinicians with a more accurate means of diagnosing CTS. The validation of a TIC will enhance the diagnostic utility of the clinical examination, thereby allowing clinicians to better select the need for additional diagnostic studies (such as EDX in the case of CTS) and the most appropriate interventions. It could also allow researchers to determine criteria for a more homogeneous patient sample for clinical trials (and thus a greater likelihood of improved outcomes with intervention studies). [58,47]

2. METHODOLOGY

2.1 Study Design

The study design is a blinded prospective diagnostic test study. A cohort of subjects was tested by one or two examiners (for reliability testing). Both the subjects and examiners were blinded to EDX test results. The examiners were also blinded to the subjects' symptom presentation.

2.2 Study Approval Process

Approval for the study was obtained from the University of Saskatchewan Behavioural Research Ethics Board on August 19th, 2004 (see Appendix D). Operational approval of the study was subsequently obtained from the Research Services Unit of the Saskatoon Health Region on September 23rd, 2004 (see Appendix E).

2.3 Subjects

The study subjects were recruited from a convenience sample of patients referred to a local Physiatry department for upper extremity EDX testing during two time periods (October-December 2004 & April-June 2005). Subjects with upper quadrant pathologies other than CTS were included in order to adequately challenge the diagnostic tests and to reflect common clinical practice. [38]

2.3.1 Inclusion Criteria:

- Patients referred to Saskatoon City Hospital Physiatry department for EDX testing due to any upper extremity complaints. Patients may or may not have been referred for suspected CTS.
- Age: 18 + years
- Informed consent given

2.3.2 **Exclusion Criteria:**

- Any shoulder, elbow or wrist pathology that limited range of movement required for safe completion of ULNT1.
- Any wrist or hand pathology that would limit safe completion of grip and pinch strength maneuvers and Phalen's test.
- Language, reading or comprehension barriers that would limit adequate completion of the study paperwork

2.4 **Measurement Procedures**

2.4.1 **Operational Definitions / Clinical Tester Preparation**

Operational definitions for all clinical tests (see Appendix F) were developed and revised prior to the data collection period in consultation with the second Physical Therapist (PT) tester. The PT testers met on two occasions prior to the first test period to review the clinical testing procedures and to ensure consistency in performance and interpretation of the tests. The testers met again prior to the second test period to review the performance and interpretation of tests. The testers had 7 and 15 years of clinical practice experience respectively and similar post graduate clinical education specializing in musculoskeletal assessment and treatment.

2.4.2 **Recruitment , Consent and Blinding**

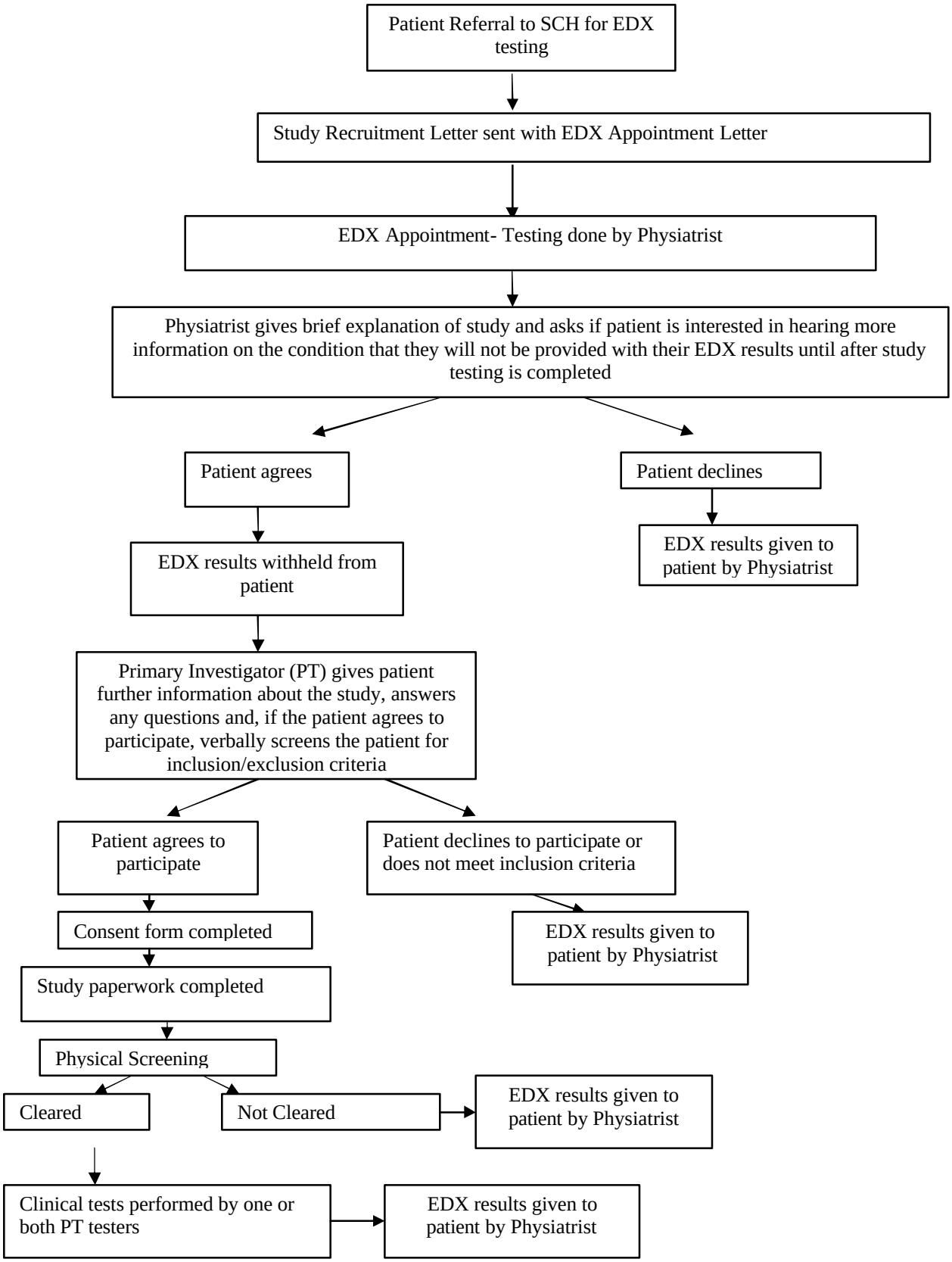
Patients referred to Saskatoon City Hospital Physiatry Department for EDX testing due to upper extremity complaints during the designated study periods were invited to participate in the study. A study recruitment letter (Appendix G) which gave a brief overview of the study was sent by the Physiatry Department along with their EDX appointment letter.

The patients attended their scheduled EDX testing appointment. Following performance of the EDX testing and prior to informing the patient of the results of the testing, the Physiatrist asked each patient if they were interested in participating in the research study. If the patient expressed an interest, then the patients were directed to talk

to the primary PT tester (Primary Investigator) for further information regarding the study. If the subject declined to participate then the results of their EDX tests were then given to them by the Physiatrist.

Interested patients were given a verbal description of the study and questions were answered. They were informed that an additional 30-60 minutes will be required in order to complete additional paperwork and clinical testing by one or two Physical Therapists. They were advised that, should they choose to participate, 10 dollars would be given to them to reimburse for their time and for parking. They were also informed that if they wished, a summary of the additional information gathered could be forwarded to their referring physician (Appendix H). Refer to Figure 2 for a diagramatic explanation of subject recruitment, consent and blinding process.

Figure 2: Subject Recruitment, Consent and Blinding Process



2.4.3 **Data Collection**

Consent forms were completed by patients who agreed to participate (Appendix I). Confidentiality of study subjects was maintained by assigning subject numbers to each subject. Only the subject's number was present on paperwork used for data collection. The key linking participant name and subject number was kept in a secure location onsite at the Physiatry Department and was destroyed upon completion of the data collection period.

Paperwork, including demographic and medical information, vocational/ avocational factors, symptom characteristics (i.e. symptom duration and location) (Appendix J), BCTQ (Appendix B) and DASH (Appendix A) questionnaires, was completed prior to performing the clinical tests.

Physical screening was done by a Physical Therapist to determine participant's ability to safely complete Phalen's, ULNT1, grip and pinch maneuvers.

Clinical measures (*Provocation tests*: Phalen's, Tinel's, carpal compression test; *Sensation testing*: Semmes Weinstein pressure sensibility, sharp/dull sensibility; *Strength testing*: thumb abduction, grip, pinch; *Neurodynamic testing*: ULNT1; *Other:* flick sign, square wrist sign, thenar eminence wasting) were performed by a Physical Therapist blinded to symptom status and to EDX test results. (See Appendix F for description and operational definitions of each test and Appendix K for clinical test reporting form.)

A second Physical Therapist tester, blinded to EDX testing results, repeated the clinical measures in the same order with a subset of study subjects to determine inter-rater reliability. The subjects that underwent repeated clinical measures were determined in a non-systematic manner and based primarily on the availability of the second PT tester.

EDX testing was done by a Physiatrist or a senior Physiatry Resident supervised by the Physiatrist (see Appendix L for overview of procedures and diagnostic criteria for CTS). The Physiatrist reported whether positive or negative EDX findings for CTS were found, the severity (mild, moderate or severe) [59]and a differential diagnosis (determined by the Physiatrist), if available (e.g. ulnar neuropathy, cervical radiculopathy), for each tested hand (see Appendix M for EDX reporting form).

3. DATA ANALYSIS

3.1 Descriptive Statistics

Descriptive statistics (mean/median, minimum, maximum, standard deviation) for continuous variables were calculated. Frequencies were determined for nominal and ordinal level variables. Characteristics (age, gender, EDX results) of patients who were excluded or who declined to participate were collected and comparisons were done (t-test for age and Chi Square for gender and EDX results) to determine if this group differed significantly from the study group.

3.2 Reliability

Inter-tester reliability of dichotomized (i.e. positive or negative) test results was determined with a kappa statistic. Inter-tester reliability of continuous measures was assessed via an intraclass correlation coefficient (ICC 2,1)[60], with corresponding standard error of measurement (SEM) and 95% confidence intervals (95% CIs). Inter-tester reliability of categorical measures was calculated by a kappa statistic and 95% CI's were calculated from the standard error (95% CI= ± (1.96)*(SE))[61]. The following qualitative interpretations for ICC's and kappa values were used in this study:

ICC Interpretation [43]:

- 'Poor'= ICC < 0.40
- 'Fair'= 0.40 = ICC < 0.70
- 'Good'= 0.70 = ICC < 0.90
- 'Excellent'= ICC = 0.90

Kappa Interpretation [47]:

- 'Poor'= Kappa < .40
- 'Fair to Good'= .40 = Kappa < .75
- 'Excellent'= Kappa = .75

3.3 **Diagnostic Validity**

Due to the lack of consensus of what constitutes the "gold standard" for CTS diagnosis, two CTS classification groups ("general" and "restricted") were formed for separate validity analysis (i.e. diagnostic test accuracy). The "CTS positive" group for the "general" classification was formed by EDX positive and probable/ possible symptom presentation *combined* with EDX negative and probable symptom presentation. The "CTS negative" group for the "general classification" was formed by EDX negative and negative symptom presentation *only.* The "restricted" classification groupings were as follows: "CTS positive" formed by EDX positive and probable or possible symptoms presentation *only* and "CTS negative" group formed by EDX negative and negative symptom presentation *combined* with EDX negative and probable symptom presentation (see Table 5).

Table 5: CTS Classification Groups

	CTS Classification Groups*	
	General	**Restricted**
Positive CTS (Symptom presentation/ EDX)	• Probable/ + or – • Possible/ +	• Probable/ + • Possible/ +
Negative (Symptom presentation/ EDX)	• Possible/ - • Unlikely/ + or -	• Probable/- • Possible/ - • Unlikely/ + or -

* Derived from Rempel 1998 [33].

The clinical examination results obtained by the first Physical Therapist examiner were used for all computations of diagnostic test accuracy.

Dichotomized tests results are required to formulate 2x2 contingency tables to calculate clinical test accuracy. The clinical test results of grip, pinch, thumb abduction strength, square wrist sign and ULNT were dichotomized into "positive" or "negative" as per the Operational Definitions (Appendix F).

Dichotomization of the self-report measures was done via receiver operator characteristic (ROC) curves. ROC curves were used to determine cutoff values for self-

report measures.[62] The cutoff value that maximized both sensitivity and specificity was selected. The symptom (BCTQ-SSS) and function (BCTQ-FN) portions of the CTQ were analyzed separately. An average score was determined for each of the BCTQ-SS and BCTQ-FN sections and a score for the DASH was determined via the method outlined in Appendix A. The corresponding cutoffs for each of BCTQ-SSS, BCTQ-FN and DASH were 2, 1.32 and 15 respectively.

The results of the questionnaires were only applied to the symptomatic side or hand. If bilateral symptoms were present, the score from the questionnaires was applied equally to both hands. If subjects had only one symptomatic side the asymptomatic hand was omitted from the validity analysis. This resulted in 13 "hands" not included in the validity analysis for the questionnaires as there were 13 subjects with a unilateral symptom presentation.

The validity of the individual items of the BCTQ-SSS was also calculated. A cutoff that was determined to be clinically relevant was used for analysis. The responses to each item are ordinal (i.e. 1 to 5) with "1" corresponding to absence of the symptom. Scores of 2 or greater were considered to be positive (i.e. the presence of the symptom was coded as "positive" and its absence was coded as "negative").

The ULNT ROM values were dichotomized using two methods. The first was if a greater than 10 degree difference in elbow ROM values was obtained from side to side, a positive classification for the more limited side was applied. The second method used was to plot an ROC curve and determine a positive cut-point that optimized both sensitivity and specificity (elbow flexion ROM = 37.5 degrees).

Dichotomization of grip and pinch strength values were determined via the method described in Appendix F. Subjects that had bilateral symptoms designated which side was "worse". The remaining side or hand was coded as "negative". Two subjects did not report which side was "worse"; however there was less than 12% difference between the grip and pinch measures on each side with one of these subjects thus each hand was coded as "negative". The remaining subject's grip and pinch data were eliminated from the validity analysis. This method of categorization resulted in 72 hands being used for the validity analysis for grip and pinch.

Sensitivity, specificity, likelihood ratios and diagnostic odds ratios along with 95% CI's [52] were calculated for each clinical test and self-report measures (DASH and CTQ) (see Appendix C for formulae). Separate analysis was performed for both "general" and "restricted" CTS classification groups. The diagnostic accuracy of individual test items was considered acceptable if LR+ was 2 or more or if LR- was 0.05 or less. [47]

3.4 **Diagnostic Test Item Cluster (TIC)**

Logistic regression was used to identify a test item cluster that may have the potential to be a better diagnostic tool than individual clinical test items. The multivariate analysis was performed separately for both CTS classification groups (i.e. general and restricted). Only those variables with acceptable accuracy (i.e. LR+ = 2 or LR- = .5) as determined by the bivariate diagnostic validity analysis were entered into the models. A forward stepwise selection procedure was used to enter variables, with P values of 0.1 for entrance to the models and 0.15 for exit from the model. The method of entry and liberal P values were chosen to prevent potentially useful variables from being excluded from the models.[58] The Hosmer-Lemeshow (HL) summary goodness of fit statistic was used to assess the fit of the models to the data and to test the hypothesis that the models fit the data (higher P values indicate a better fit).[63]

Variables selected by the regression models as diagnostic of CTS were used to determine various potentially valid TICs. The first strategy employed was to use the logistic regression strictly as a selection method to identify test items that are likely to be most diagnostic for CTS. The test items identified were combined into various "test positive" levels of the TIC. For example, if items A, B and C were chosen the following three "test positive" levels would be considered:

1. at least one of A, B or C were positive (TIC 1)
2. at least 2 of A, B or C were positive (TIC 2)
3. all three of A, B and C were positive (TIC 3)

The second strategy was to use the items selected by the logistic regression and combine them into a single "adjusted" TIC. The sum of the products of the results of each test (X) with the corresponding β coefficients was used to determine the "adjusted" TIC (i.e. β_0 +

$\beta_A X_A + \beta_B X_B + \beta_C X_C$ = "adjusted" TIC). The "adjusted" TIC, a continuous variable, was subsequently dichotomized via plotting an ROC curve to identify the most appropriate cut-point for a positive test (i.e. the value that related to the highest sensitivity and specificity).

The diagnostic validity (i.e. sensitivity, specificity, LRs and DORs) was calculated for each of the identified TICs as previously described for other dichotomous variables.

4. **RESULTS**

4.1 **Descriptive Statistics**

A total of 60 patients were eligible to participate in the study during the data collection time periods. Twenty patients declined to participate and three were excluded prior to data collection by the Primary Investigator (2 had insufficient wrist range of motion due to recent colles fractures and one was unable to read and comprehend the paperwork). The non-participant group characteristics of age, gender and EDX results were collected and compared to the study participants. No significant differences were found between the groups for age (p=.81), gender (p=.22) or EDX results (p=.75).

The descriptive statistics for age and duration of symptoms of the 37 participants in the study are listed in Table 6. The descriptive statistics for other demographic, vocational and medical information are listed in Table 7. Table 8 presents the descriptive statistics per hand (n=74) for the CTS classification, EDX results, symptom location (according to hand diagram) and symptom side.

The prevalence of CTS according to the "general" classification was 66.2% (49/74) and the prevalence according to the "restricted" classification was 51.4% (38/74).

Table 6: Descriptive Statistics for Continuous Variables

Variable	Valid N	Mean	Minimum	Maximum	SD
Age (years)	37	50.86	19	85	16.64
Symptom Duration (months)	34	33.26	3	360	62.36

Table 7: Descriptive Statistics for Categorical Variables

Variable	Frequency	Percent
Gender		
Female	25	67.6
Male	12	32.4
Marital		
Married	25	67.6
Divorced/Separated/ Widowed	7	18.9
Never Married	5	13.5
Education		
< Grade 12	8	21.6
Grade 12	14	37.8
University / Trade School	15	40.5
Income		
<15,000	6	16.2
16-29000	6	16.2
30-59000	18	48.6
60-99000	5	13.5
>100,000	1	2.7
BMI* [64]		
<20 kg/m^2 (less than desirable)	1	2.7
20-24.9 kg/m^2 (desirable)	7	18.9
25-29.9 kg/m^2 (grade 1 obesity)	13	35.1
30-40 kg/m^2 (grade 2 obesity)	16	43.2
Hand Dominance		
Right	30	81.1
Left	6	16.2
Both	1	2.7
Work		
Full Time	21	56.8
Part Time	3	8.1
Unemployed	1	2.7
House Work	7	18.9
Student	2	5.4
Retired	3	8.1
Repetitive Work	23	62.2
Cold Work	10	27.0
Vibration Work	8	21.6
WCB**	2	5.4
SGI***	0	0
Diabetes****	6	16.2
Rheumatoid Arthritis****	3	8.1
Osteoarthritis****	3	8.1
Wrist fracture****	1	2.7
Thyroid dysfunction****	7	18.9
Smoker****	6	16.2

*BMI= Body Mass Index

**WCB= Worker's Compensation Board

***SGI= Saskatchewan Government Insurance (i.e. post motor vehicle accidents)
****Subject self-report as per completed "General Paperwork" Appendix I

Table 8: EDX Results, Symptom Characteristics and Location

Variable	Frequency	Percent
CTS Classification:		
Negative	25	33.8
Classic/ Negative EDX	11	14.9
Possible/Positive EDX	4	5.4
Classic/Positive EDX	34	45.9
EDX:		
Negative	26	35.1
Mild	14	18.9
Moderate	26	35.9
Severe	8	10.8
Hand Diagram:		
Unlikely	23	31.1
Possible	5	6.8
Probable	46	62.2
Symptom side		
Right	6	16.2
Left	6	16.2
Both	25	67.2
Worse side		
Right	15	40.5
Left	10	27.0

4.2 **Reliability**

Inter-tester reliability was computed using the results from the paired results of the first and second PT testers. There were 24 subjects (48 hands) that were examined by both testers. The 13 subjects not included in the reliability analysis did not differ from the other 24 subjects with regard to age (p=.72) and gender (p=.93).

The reliability coefficients for the clinical examination items and their associated 95% CI's are listed in Table 9. Eight items had kappa values at least fair or better (kappa =0.40). One item (flick sign) had an "excellent" kappa value (=0.75). Three items had ICC values that were "excellent" (ICC=0.90).

Table 9: Inter-tester Reliability of CTS Clinical Measures

Clinical Test	Kappa	ICC (2,1)	SD (pooled)	SEM***	95 % CI ****
Phalen's*	.58	--	--	--	(.35-.81)
Tinel's*	.51	--	--	--	(.15-.87)
Carpal Compression*	.67	_	_	_	(.46-.88)
Sharp/Dull Sensibility*	.70	_	_	_	(.50-.90)
Pressure Sensibility*	.64	_	_	_	(.39-.89)
Grip Strength**	**_**	**.99**	**26.43 kg**	**3.01 kg**	**(.98-.99)**
Tip to Tip Pinch**	_	.76	2.98 kg	1.47 kg	(.57-.86)
Key Pinch**	**_**	**.92**	**4.98 kg**	**1.40 kg**	**(.86-.96)**
3 Jaw Chuck Pinch**	**_**	**.93**	**4.32 kg**	**1.13 kg**	**(.88-.96)**
Thumb Abduction Strength*	.18	_	_	_	(-.04-.40)
Flick Sign*	**.75**	**--**	**--**	**--**	**(.48-.98)**
Thenar Atrophy*	-.08	_	_	_	(-.14-.02)
Square Wrist Ratio**	_	.11	5.75 mm	5.42mm	(-.59-.50)
Square Wrist Ratio*	.05	_	_	_	(-.22-.32)
ULNT 1 ROM**	_	.66	22.76^{0}	13.29^{0}	(.39-.81)
ULNT 1 symptom provocation*	.70	_	_	_	(.50-.90)
ULNT 1 >10deg diff*	.05	_	_	_	(-.24-.34)
ULNT 1 SF provocation*	.49	_	_	_	(.24-.74)

* Categorical Measures
** Continuous Measures
***SEM Calculated as a function of the pooled group SD and the ICC (SEM= SD(pooled) x $\sqrt{}$ (1- ICC)) [43]
**** 95% CIs for ICC's determined by SPSS V10 softwear. 95% CIs for Kappa values calculated from SEM provided by SPSS (± 1.96* SEM).

4.3 **Diagnostic Validity**

The sensitivity, specificity, likelihood ratios and diagnostic odds ratios for each variable and their associated 95% CIs are listed in Tables 10 and 11. The following variables were found to have acceptable diagnostic validity when a "general" CTS classification was used: Phalen's, carpal compression test, grip strength, thumb abduction strength, flick sign, thenar atrophy, ULNT symptoms, BCTQ-FN, item 1 (i.e. hand or wrist pain at night) and item 6 (i.e. numbness in hand) of the BCTQ-SSS. The following variables were found to have acceptable diagnostic validity when a "restricted" CTS classification was used: flick sign, thenar atrophy, ULNT symptoms, BCTQ-SSS, and item 6 (i.e. hand numbness) and item 11 (i.e. difficulty with grasping and use of small objects) of the BCTQ-SSS.

Table 10: Univariate Validity Analysis Summary- Clinical Tests

Test Item	CTS Classification	Sensitivity % (95%CI)	Specificity % (95%CI)	+LR (95%CI)	-LR (95%CI)	Diagnostic OR (95%CI)
Phalen's	General*	53.06 (39.38,66.30)	84.00 (65.35,93.60)	**3.32 (1.30,8.46)**	.56 (.40,.79)	5.94 (1.77,19.85)
	Restricted**	52.63 (37.26,67.52)	72.22 (56.01,84.15)	1.90 (1.03,3.48)	.66 (.44,.97)	2.89 (1.10, 7.61)
Tinel's	General*	12.24 (5.73,24.24)	92.00 (75.03,97.78)	1.53 (.33,7.04)	.95 (.82, 1.12)	1.61 (.30, 8.60)
	Restricted**	13.16 (5.75,27.33)	91.67 (78.17,97.13)	1.58 (.41,6.13)	.95 (.81,1.11)	1.67 (.37, 7.55)
CCT	General*	55.10 (41.32,68.15)	80.00 (60.87, 91.14)	**2.76 (1.21, 6.28)**	.56 (.39, .81)	4.91 (1.59, 15.20)
	Restricted**	52.63 (37.26,67.52)	66.67 (50.33, 79.79)	1.58 (.91, 2.74)	.71 (.47,1.07)	2.22 (.87, 5.70)
Sharp/ Dull	General*	40.82 (28.22, 54.75)	72.00 (52.42, 85.72)	1.46 (.71, 2.98)	.82 (.59, 1.15)	1.77 (.63, 5.03)
	Restricted**	36.84 (23.38, 52.72)	63.89 (47.58, 77.52)	1.02 (.56, 1.86)	.99 (.7, 1.40)	1.03 (.40, 2.66)
Pressure Sensation	General*	14.29 (7.1, 26.67)	88.00 (70.04, 95.83)	1.19 (.34, 4.21)	.97 (.81, 1.17)	1.22 (.29, 5.20)
	Restricted**	13.16 (5.75, 27.33)	86.11 (71.34, 93.92)	.95 (.30, 3.00)	1.09 (.84, 1.21)	.94 (.25, 3.56)
Grip Strength	General*	21.28 (11.99,34.90)	92.00 (75.03,97.78)	**2.66 (.63,11.21)**	.86 (.71,1.03)	3.11 (.62,15.47)
	Restricted**	13.51 (5.91,27.98)	80.00 (64.11,89.96)	.68 (.24,1.93)	1.08 (.88,1.33)	.63 (.18,2.19)
Tip to tip pinch	General*	19.15 (10.42,32.54)	84.00 (65.35,93.60)	1.20 (.41, 3.50)	.96 (.77,1.20)	1.24 (.34, 4.15)
	Restricted**	13.51 (5.91, 27.90)	77.14 (60.98,87.93)	.59 (.21,1.64)	1.12 (.90,1.40)	.53 (.15,1.80)
Key Pinch	General*	20.00 (10.90,33.82)	88.00 (70.04, 95.83)	1.67 (.50,5.60)	.91 (.74,1.12)	1.83 (.45,7.51)
	Restricted**	14.29 (6.26,29.38)	80.00 (64.11,89.96)	.71 (.25,2.04)	1.07 (.87,1.33)	.67 (.19,2.35)
3 Jaw Chuck Pinch	General*	19.15 (10.42,32.54)	84.00 (65.35,93.60)	1.20 (.41,3.50)	.96 (.77,1.20)	1.24 (.34,4.53)
	Restricted**	16.22 (7.65,31.14)	80.00 (64.11,89.96)	.81 (.30,2.18)	1.05 (.84,1.30)	.77 (.23,2.58)
Thumb Abduction	General*	34.69 (22.92, 48.69)	84.00 (65.35, 93.60)	**2.17 (.82, 5.76)**	.78 (.60, 1.02)	2.79 (.82, 9.45)
	Restricted**	36.84 (23.38, 52.72)	80.56 (64.97, 90.25)	1.90 (.87, 4.15)	.78 (.59, 1.05)	2.42 (.84, 6.95)
Flick sign	General*	30.61 (19.52, 44.53)	96.00 (80.47, 99.29)	**7.65 (1.07,54.66)**	.72 (.59, .89)	10.59 (1.31, 85.66)
	Restricted**	36.84 (23.38, 52.72)	94.44 (81.86, 98.46)	**6.63 (1.62,27.16)**	.67 (.52, .86)	9.92 (2.06, 47.72)
Thenar Atrophy	General*	12.24 (5.73, 24.24)	98.04 (83.70, 99.79)	**6.25 (.36,107.42)**	.90 (.80,1.01)	6.98 (.37, 130.24)
	Restricted**	13.16 (5.75, 27.33)	97.22 (85.83, 99.51)	**4.74 (.58, 38.61)**	.89 (.78, 1.02)	5.30 (.59, 47.82)
Square Wrist Sign	General*	34.69 (22.92, 48.69)	76.00 (56.57, 8.85)	1.45 (.65, 3.21)	.86 (.64, 1.16)	1.68 (.57, 5.00)

Table 10: Univariate Validity Analysis Summary- Clinical Tests (continued)

Square Wrist Sign	Restricted**	39.47 (25.60, 55.28)	77.78 (61.92, 88.28)	1.78 (.86, 3.68)	.78 (.57, 1.06)	2.28 (.82, 6.33)
ULNT Symptoms	General*	40.82 (28.22, 54.75)	80.00 (60.87, 91.14)	**2.04 (.87, 4.79)**	.74 (.55, 1.00)	2.76 (.89, 8.57)
	Restricted**	44.74 (30.15, 60.29)	77.78 (61.92, 88.28)	**2.01 (.99, 4.08)**	.71 (.51, .99)	2.83 (1.03, 7.80)
ULNT ROM 10 degree diff	General*	14.29 (7.1, 26.67)	84.00 (65.35, 9.36)	.89 (.29, 2.76)	1.02 .83, 1.25)	.88 (.23, 3.33)
	Restricted**	7.89 (2.72, 20.80)	77.78 (61.92, 88.28)	.36 (.10, 1.24)	1.18 (.97, 1.44)	.30 (.07, 1.24)
ULNT ROM categorical (ROC curve)	General*	4.08 (1.13, 13.71)	92.00 (75.03, 97.78)	.51 (.08, 3.41)	1.04 .91, 1.19)	.49 (.07, 3.70)
	Restricted**	5.26 (1.46, 17.29)	94.44 (81.86, 98.40)	.95 (.14, 6.37)	1.00 (.90, 1.12)	.94 (.13, 7.09)
ULNT SF	General*	40.82 (28.22, 54.75)	76.00 (56.57, .89)	1.70 (.78, 3.69)	.78 (.57, 1.07)	2.18 (.74, 6.43)
	Restricted**	42.11 (27.85, 57.81)	72.22 (56.01, 84.15)	1.52 (.80, 2.89)	.80 (.57, 1.13)	1.89 (.72, 5.00)
BCTQ-SSS	General*	74.47 (60.49, 84.75)	50.00 (26.80, 73.20)	1.49 (.86, 2.58)	.51 (.25, 1.05)	2.92 (.85, 10.04)
	Restricted **	81.08 (65.80, 90.52)	50.00 (31.43,68.57)	1.62 (1.06, 2.50)	**.38 (.17, .82)**	4.29 (1.36, 13.50)
BCTQ- FN	General*	76.60 (62.78, 86.40)	64.29 (38.78,83.66)	**2.15 (1.04, 4.41)**	**.37 (.19, .70)**	5.89 (1.63, 21.29)
	Restricted**	75.68 (59.88, 86.64)	45.83 (27.89,64.93)	1.40 (.93, 2.11)	.53 (.26, 1.09)	2.63 (.88, 7.90)
DASH	General*	61.07 (47.43, 74.21)	42.86 (21.38,67.41)	1.08 (.65, 1.79)	.89 (.44, 1.81)	1.21 (.36, 4.06)
	Restricted**	62.16 (46.10,75.94)	41.67 (24.47,61.17)	1.07 (.70,1.62)	.91 (.49,1.70)	1.17 (.41,3.35)

***** "CTS positive" group formed by EDX positive and probable or possible symptoms presentation *combined* with EDX negative and probable symptom presentation. "CTS negative" group formed by EDX negative and negative symptom presentation *only.*

****** "CTS positive" group formed by EDX positive and probable or possible symptoms presentation *only.* "CTS negative" group formed by EDX negative and negative symptom presentation *combined* with EDX negative and probable symptom presentation.

Clinically significant items are in bold face type.

Table 11: Univariate Validity Analysis Summary- BCTQ- SSS

Test Item	CTS Classification	Sensitivity % (95%CI)	Specificity % (95%CI)	+LR (95%CI)	-LR (95%CI)	Diagnostic OR (95%CI)
SSS 1	General*	72.34 (58.24, 83.06)	57.14 (32.59, 78.62)	1.69 (.90,3.17)	**.48 (.25, .93)**	3.49 (1.01, 12.00)
	Restricted**	72.97 (57.02, 84.60)	45.83 (27.89, 64.93)	1.35 (.89,2.04)	.59 (.30, 1.17	2.29 (.77, 6.74)
SSS 2	General*	59.57 (45.34, 72.36)	35.71 (16.34, 61.24)	.93 (.59,1.46)	1.13 (.52, 2.48)	.82 (.24, 2.83)
	Restricted**	62.16 (46.10, 75.94)	41.67 (24.47, 61.17)	1.07 (.70,1.62)	.91 (.49, 1.70)	1.17 (.41, 3.35)
SSS 3	General*	70.21 (56.02, 81.35)	21.43 (7.57, 47.59)	.89 (.64,1.24)	1.39 (.47, 4.16)	.64 (.16, 2.66)
	Restricted**	70.27 (54.22, 82.51)	25.00 (12.00, 44.90)	.94 (.69,1.28)	1.19 (.51, 2.79)	.79 (.25, 2.50)
SSS 4	General*	65.96 (51.67, 77.83)	35.71 (16.34, 61.24)	1.03 .66, 1.60)	.95 (.43, 2.14)	1.08 (.31, 3.75)
	Restricted**	64.86 (48.76, 78.17)	33.33 (17.97, 53.29)	.97 (.67, .1.41)	1.05 (.51, 2.16)	.92 (.31, 2.73)
SSS 5	General*	65.96 (51.67, 77.83)	21.43 (7.57, 47.59)	.84 (.60, 1.18)	1.60 (.54, 4.67)	.53 (.13, 2.17)
	Restricted**	70.27 (54.22, 82.51)	52.94 (36.74, 68.55)	1.49 (.99, 2.26)	.56 (.31, 1.01)	2.66 (1.00, 7.05)
SSS 6	General*	95.74 (85.75, 98.83)	35.71 (16.34, 61.24)	1.49 (1.00,2.21)	**.12 (.03, .55)**	12.50 (2.09, 74.81)
	Restricted**	97.30 (86.18, 99.52)	25.00 (.12, .45)	1.30 (1.02,1.64)	**.11 (.01, .84)**	12.00 (1.34,107.36)
SSS 7	General*	69.23 (55.73, 80.09)	57.14 (32.59, 78.62)	1.62 (.86, 3.04)	.54 (.29, .99)	3.00 (.89, 10.07)
	Restricted**	67.57 (51.46, 80.37)	50.00 (31.43, 68.57)	1.35 (.86, 2.14)	.65 (.35, 1.20)	2.08 (.73, 5.99)
SSS 8	General*	89.19 (80.07, 96.64)	14.29 (4.01, 39.94)	1.07 (.85, 1.35)	.60 (.12, 2.92)	1.79 (.29, 10.99)
	Restricted**	89.19 (75.29, 95.71)	8.33 (2.32, 25.85)	.97 (.83, 1.15)	1.30 (.26, 6.54	.75 (.13, 4.45)
SSS 9	General*	87.23 (74.83, 94.02)	14.29 (4.01, 39.94)	1.02 (.80, 1.29)	.89 (.20, 3.95)	1.14 (.20, 6.39)
	Restricted**	86.49 (72.02, 94.09)	12.50 (4.34, 31.00)	.99 (.81, 1.20)	1.08 (.28, 4.11)	.91 (.20, 4.24)
SSS 10	General*	72.34 (58.24, 83.06)	35.71 (16.34, 61.24)	1.13 (.73, 1.73)	.77 (.33, 1.80)	1.45 (.41, 5.16)
	Restricted**	75.68 (59.88, 86.64)	37.50 (21.16, 57.29)	1.21 (.85, 1.74)	.65 (.30, 1.40)	1.87 (.61, 5.70)
SSS 11	General*	72.34 (58.24, 83.06)	42.86 (21.38, 67.41)	1.27 (.78, 2.06)	.65 (.30, 1.38)	1.96 (.57, 6.76)
	Restricted**	78.38 (62.80, 88.61)	45.83 (27.89, 64.93)	1.45 (.97, 2.17)	**.47 (.22, 1.00)**	3.07 (1.00, 9.41)

Clinically significant items (i.e. +LR = 2 or –LR = 0.5) are in bold face type.

4.4 **Diagnostic Test Item Cluster**

Separate analyses for general and specific CTS classification were performed. The 10 variables that were found to have acceptable diagnostic accuracy for the "general" CTS classification were entered into a regression model as potential predictors for CTS. (Refer to Appendix N for SPSS logistic regression results tables). After list-wise deletion, a total of 59 "hands" were used in the analysis (45 hands with CTS and 14 control hands). The results of the HL test indicated that the model fit the data (p=1.00). The following two variables were chosen by the model: BCTQ-SSS item 6 (presence of numbness as a symptom) and ULNT1 symptoms (i.e. reproduction of all or part of subject's symptoms with ULNT1 test). Two "test positive" conditions of the TIC were examined for diagnostic validity: TIC 1 (at least 1 of the 2 identified items positive) and TIC 2 (both of the identified items positive). An "adjusted" TIC was also determined based on the β coefficients from the logistic regression analysis. The following formula for the "adjusted" TIC model was used: $\beta_0 + \beta_A X_A + \beta_B X_B$. Where β_0=-22.83, β_A=23.97, X_A = BCTQ-SSS item 6, β_B= 1.86, and X_B=ULNT1 symptoms. A continuous value for the adjusted TIC was derived from this formula so an ROC curve was plotted (using SPSS Version 10) to determine the optimal cut-point (-9.92) for a "positive" adjusted TIC. The adjusted TIC was subsequently dichotomized based on this cut-point. The adjusted TIC yielded the same results as the single best test item (SSS item 6).

The 6 variables that were found to have acceptable diagnostic accuracy for the "restricted" CTS classification were also entered into a regression model as potential predictors of CTS. After a list-wise deletion of 13 hands (BCTQ items n=61), a total of 61 hands were used in the analysis (37 hands with CTS and 24 control hands). The results of the HL test indicated that the model fit the data (p=.74). The following two variables were chosen by the model: BCTQ-SSS item 6 and BCTQ-SSS (average score =2 from BCTQ-SSS). Two "test positive" conditions of the TIC were examined for diagnostic validity (TIC 1 and TIC 2). The following formula for the "adjusted" TIC model was used: $\beta_0 + \beta_A X_A + \beta_B X_B$. Where β_0=-2.23, β_A=2.09, X_A = CTQ-SS item 6, β_B= 1.17, and X_B=BCTQ-SSS. A continuous value for the adjusted TIC was derived from this formula so an ROC curve was plotted (using SPSS Version 10) to determine the optimal cut-point (.45) for a "positive" adjusted TIC. The adjusted TIC was subsequently

dichotomized based on this cut-point. Based on this cut point the adjusted TIC yielded the same results as the TIC 2.

Table 12 provides a summary of the diagnostic validity of TIC 1, TIC 2, and the adjusted TIC and the single best test item from the bivariate analysis (BCTQ-SSS item 6) for both the general and restricted classification schemes.

Table 12: Diagnostic Validity- Multivariate Analysis vs. Single Best Test Item

Classification	Test Item	Sensitivity % (95%CI)	Specificity % (95%CI)	+LR (95%CI)	-LR (95%CI)	Diagnostic OR (95%CI)
General	TIC 1	95.74 % (85.75,98.83)	7.14% (1.27, 31.47)	1.03 (.88,1.21)	.60 (.06, 6.09)	1.73 (.15, 20.64)
	TIC 2	42.55% (29.51,56.72)	92.86% (68.53, 98.73)	5.96 (.88, 40.54)	.62 (.47, .82)	9.63 (1.16, 79.79)
	Adjusted TIC	95.74% (85.75, 98.83)	35.71% (16.34, 61.24)	1.49 (1.00, 2.21)	.12 (.03, .55)	12.5 (2.09, 74.81)
	Single Best Test Item: SSS Item 6	95.74% (85.75, 98.83)	35.71% (16.34, 61.24)	1.49 (1.00, 2.21)	.12 (.03, .55)	12.5 (2.09, 74.81)
Restricted	TIC 1	97.30% (86.18, 99.52)	16.67% (6.68,35.85)	1.17 (.97,1.41)	.16 (.02,1.37)	7.20 (.75,68.89)
	TIC2	81.08% (65.80,90.52)	58.83% (38.83,75.53)	1.95 (1.18,3.20)	.32 (.15,.69)	6.00 (1.89,19.06)
	Adjusted TIC	81.08% (65.80,90.52)	58.83% (38.83,75.53)	1.95 (1.18,3.20)	.32 (.15,.69)	6.00 (1.89,19.06)
	Single Best Test Item: SSS item 6	97.30 % (86.18, 99.52)	25.00 % (.12, .45)	1.30 (1.02,1.64)	.11 (.01, .84)	12.00 (1.34,107.36)

4.5 **Summary of Results**

A total of 37 subjects participated in this study resulting in 74 "hands" used as the primary unit of analysis. The categorical clinical examination items that had kappa values of .40 (i.e. fair) or greater were: Phalen's (.58), Tinel's (.51), carpal compression (.67), sharp/ dull sensibility(.70), pressure sensibility (.64), flick sign (.75), ULNT1 symptom provocation (.70), and ULNT1 SF provocation (.49). The following variables were found to have acceptable diagnostic validity when a "general" CTS classification was used: Phalen's (+LR=3.32), carpal compression test (+LR=2.76), grip strength (+LR=2.66), thumb abduction strength (+LR=2.17), flick sign (+LR=7.65), thenar atrophy (+LR=6.25), ULNT symptoms (+LR=2.04), BCTQ-FN (+LR=2.15, -

LR=.37), item 1 (i.e. hand or wrist pain at night) (-LR=.48) and item 6 (i.e. numbness in hand) (-LR=.12) of the BCTQ-SSS. The following variables were found to have acceptable diagnostic validity when a "restricted" CTS classification was used: flick sign (+LR=6.63), thenar atrophy (+LR=4.74), ULNT symptoms (+LR=2.01), BCTQ-SSS (-LR=.38), and item 6 (i.e. hand numbness) (-LR=.11) and item 11 (i.e. difficulty with grasping and use of small objects) (-LR=.47) of the BCTQ-SSS. The TIC derived for the general classification grouping was item 6 of the SSS and ULNT 1. The TIC derived for the restricted classification was item 6 of the SSS and overall score of the BCTQ-SSS.

5. **DISCUSSION**

5.1 **Study Design and Potential for Bias**

The optimal design for any research study is the one that most effectively reduces susceptibility to bias. The determination of the validity of diagnostic tests is best accomplished through "a prospective, blind comparison of the test and the reference test in a consecutive series of patients from a relevant clinical population."[54] Beyond the basic design there is potential for bias through other study variables such as the reference standard, the diagnostic test and the study population.

The reference standard used in this study was the combination of EDX testing and symptom location (as per the hand diagram Appendix J). The extent to which this reference standard determines the "true" presence of CTS is debatable[33], although the use of both "general" and "restricted" classification criteria in this study does allow for a broader interpretation of the results in the absence of a "perfect" gold standard.

The reference standard used should also be consistent with the intended purpose of the diagnostic test[47]. The use of EDX test results combined with symptom location is a pathoanatomical reflection of the presence median nerve pathology at the wrist and hand (i.e. CTS diagnosis); the intended use of the clinical tests and self-report items in this study was to identify the presence of CTS (i.e. diagnosis of CTS).

Verification bias occurs when not all subjects are assessed by use of the reference standard in the same way. [65,66] The EDX testing was primarily done by the same Physiatrist using the outlined criteria (Appendix L). A Physiatry resident who used the same diagnostic criteria and who was supervised by the Physiatrist also performed some of the testing. The symptom location was determined via the completed hand diagram (Appendix J). The symptom pattern (i.e. probable, possible or unlikely) was classified by the primary investigator after the clinical tests had been performed. The reference standard used (i.e. EDX results and symptom pattern) was consistently applied to all subjects, thus verification bias was minimized.

The reference standard should be independent of the diagnostic test to avoid incorporation bias which is likely to inflate the accuracy of a diagnostic test[54]. EDX testing is independent from any of the clinical tests and self-report items. The hand

diagram was used to determine symptom pattern; however, it could be construed as having some of the same elements as the symptom portion of the BCTQ. Independence of these elements was likely maintained as the classification via the hand diagram was done solely on the basis of symptom location and not on the basis of the type (i.e. pain, numbness, tingling) and severity of the symptoms which were quantified by the BCTQ.

Review bias may occur if either the reference standard or the diagnostic test is judged by an individual with knowledge of the other result or knowledge of the overall clinical presentation of the subject.[54,67] If blinding is not maintained, judgment of either the reference standard or the clinical tests may be influenced by expectations based on knowledge of the test results or other clinical information. The Physiatrist and resident were blinded to the results of the clinical tests and self-report items as these occurred after the EDX testing was completed. They were not necessarily blinded to the overall clinical presentation of the subject as a standard history of symptoms was performed prior to testing. Thus, judgment of the EDX results may have been somewhat influenced by the expectations arising out of this additional clinical and historical knowledge. The Physical Therapists performing the clinical tests as well as the participants were blinded from the EDX results until after completion of the study paperwork (including self-report items) and the clinical testing. Blinding the participants from the EDX results was essential to prevent that knowledge from influencing their responses to the self-report questionnaires or the clinical tests. The classification of symptom location was done by the primary investigator following completion of the data collection period of the study. The hand diagram results were separated from the EDX results prior to data entry and no attempts to link EDX results and hand diagram classification was done at that time.

A diagnostic test study should include subjects who would be likely to undergo the test(s) in clinical practice. [54,67] Spectrum (or selection) bias occurs when study subjects are not representative of the population on whom the test is typically applied in practice.[54] Avoidance of spectrum bias can be achieved by utilizing a prospective cohort design. [67] The subjects for this study were drawn prospectively from a consecutive group of patients from a clinical population. The descriptive characteristics of the study sample (Table 5) appear to reflect previous research examining associated factors for the development of CTS (refer to "CTS Associated Factors" p. 2).

Individuals who are positive on the reference standard should also reflect a continuum of severity from mild to severe, and individuals who have a negative result on the reference standard should have conditions commonly confused with the condition of interest.[68] Positive EDX results were classified for each hand as "mild", "moderate" or "severe" by the Physiatrist (Table 7). A range of values in all categories was obtained with most (35.9%) being classified as "moderate". The study Physiatrist identified the following differential diagnoses among subjects: osteoarthritis of the first carpo-metacarpal (thumb) joint, trigger finger, neck/shoulder/soft tissue pain, ulnar neuropathy, radial neuropathy, diabetic neuropathy, generalized neuropathy, tennis elbow, and focal dystonia. Most of these are conditions that may be commonly confused with CTS.

5.2 **The Role of Reliability**

This study attempted to quantify inter-rater reliability among clinical diagnostic tests for CTS. Beyond the numerous potential sources of variability between the two testers, there may have been variability arising from any instruments used (i.e. pinch and grip dynamometers, monofilaments, calipers and goniometer) or among each subject from one testing period to the other. A partially standardized approach was used whereby some aspects of potential variability were controlled through the use of pre-determined operational definitions of the clinical tests, indicating method of application and criteria for positivity (Appendix F), and two training sessions for the testers ensured similarity in application and interpretation of the clinical tests. This partially standardized approach tends to be more clinically useful than a highly standardized approach, such as in a lab environment, and yields better reliability than a non-standardized approach.[69]

Comparison of the reliability of the clinical tests examined in this study to prior research shows some variability. The inter-tester reliability of Phalen's and Tinel's tests had previously been reported by Kulhman & Hennessey[42] to have kappa values of .65 and .79 respectively; confidence intervals were not reported in this cited study. These point estimates for reliability differ from this study's findings (Phalen's kappa=.58 Tinel's kappa= .51); however the 95% confidence intervals (Phalen's 95%CI= .38-.81, and Tinel's 95% CI= .15-.87) are inclusive of the kappa values previously determined by

Kulhman & Hennessey.[42] The reliability of the carpal compression test, sharp/dull sensibility, thumb abductor strength test, flick sign, thenar atrophy and square shaped wrist had previously not been reported. Among these the only ones to demonstrate a kappa value corresponding to "poor" reliability were thumb abduction strength, thenar atrophy and square wrist sign. The subjective interpretation of thumb abduction strength and thenar atrophy was a likely source of variability between testers. Variability of the square wrist sign was potentially influenced by measurement error in either the application of the calipers, the degree of pressure applied, or the reading of the caliper measurement. Grip strength, key pinch and 3 jaw tripod pinch showed excellent ICC values (i.e. >.90). These values are comparable to previous research (Table 9). These high values are surprising given the many potential sources of variability with these tests such as instrument variation, reading errors, variable subject positioning, influences of subject pain, and subject effort. The lower reliability of tip to tip pinch (ICC=.76) is likely a reflection on the variability of subject positioning. Despite instructions provided by the testers, subjects in the study often adopted more of a "key pinch" positioning with pressure applied though the radial aspect of the index finger instead of through the palmar aspect.

Figure 3: Tip to Tip Pinch vs Key Pinch

Among the reliability of the various categorical components of the ULNT (i.e. symptom provocation, ROM>10 degrees difference, and SF provocation), symptom provocation demonstrated the highest kappa value (k=.70, 95%CI .50-.90). The continuous variable of elbow ROM had an ICC of .66 (95%CI .39-.81). This is in stark contrast to Coppieter et al.'s[43] study in which a reliability value of ICC=.98 for ULNT ROM was obtained. Equally high levels of inter-tester reliability for both a highly

standardized and a partially standardized more clinical approach were reported. The protocol for both involved an electrogoniometer attached to the medial side of the elbow to measure ROM. This study utilized a regular goniometer that was applied to the medial aspect of the elbow with fabricated velcro straps. Maintaining the alignment of the fixed and moving arms of the goniometer in order to obtain an accurate reading was difficult, if not impossible, for one tester alone to do while still attempting to maintain the upper extremity positioning and scapular stabilization required for performance of the ULNT (Figure 4). The use of a regular goniometer in this manner is, however, demonstrative of how the test may be carried out in a clinical scenario in the absence of the availability of an electrogoniometer.

Figure 4: Goniometer Placement with ULNT1

Coppieters et al. [43] do not clarify how the electrogoniometer was affixed or if an assistant was used for the test procedure; these are factors that may potentially influence test reliability.

Optimization of the reliability of a test may result in improved diagnostic validity.[70] The reliability of the ROM component of the ULNT demonstrated poor reliability and validity (Tables 8 & 9) within this study; however, perhaps performing the test in a more reliable manner would result in improved diagnostic validity.[69]

Reliability has traditionally been emphasized as a precursor to validity; the numerous studies examining test reliability without any assessment of validity attest to this approach. Unfortunately a risk of this is that it may lead to the dismissal of potentially useful tests based on the inability to reach an arbitrary threshold of reliability.[47] Also reliability assessments conducted separately from an examination of validity may result in the promotion of highly reliable but diagnostically meaningless

tests.[47] For these reasons reliability data should be considered a compliment to and not only a pre-requisite of an assessment of diagnostic value.

5.3 **Validity**

Prior studies examining the diagnostic validity of clinical tests for CTS have, for the most part, reported a wide range of point estimates of validity (Table 3), thus comparison to the values obtained in this study is somewhat difficult.

The sensitivity, specificity and LR values in this study are taken from a sample and represent an estimate of the true value that could be found in the population[52]. The confidence interval (CI) indicates the precision of this estimate. In this case the CI's for all values obtained are quite wide and many include values that are not clinically valid (i.e. sensitivity or specificity= 50% or LR=1), thus the usefulness of the tests may be questionable. The width of the CI is often related to the sample size; a priori sample size calculation was not carried out in this study. The clinical tests and self-report items that were found to be *both* clinically and statistically valid were (G=general classification, R=restricted classification): Phalen's (G), carpal compression test (G), Flick sign (G&R), BCTQ-Symptom Severity Scale (R), BCTQ-Function (G), SSS 1 (G), and SSS 6 (G & R).

The diagnostic odds ratio (DOR) values for the clinical tests obtained with the general CTS classification were generally higher than the restricted classification. This is expected as the general classification is more inclusive of what constitutes a positive test. The exceptions to this were Tinel's (G=1.61, R=1.67), square wrist sign (G=1.68, R=2.28), and ULNT symptoms (G=2.76, R=2.83), ULNT ROM (G=.49, R=.94). Among the self-report measures, the overall BCTQ-SSS score (G=2.92, R=4.29), and items 2(G=.82, R=1.17), 3 (G=.64, R=.79), 5 (G=.53, R=2.66), 10 (G=1.45, R=1.87) and 11 (G=1.96, R=3.07) of the BCTQ-SSS demonstrated higher DOR's for the "restricted" than the "general" classification groups. Conversely, the functional component of the BCTQ demonstrated higher DOR's for the "general" than the "restricted" classification groups (G=5.89, R=2.63).

Levine et al[46] developed the BCTQ-SS to be a reflection of the severity of symptoms in CTS. The tool was developed through consultation with medical experts

and people with CTS in order to develop domains for the symptom severity scale. You et al.[71] found that the BCTQ-SSS had a "significant relationship" to nerve conduction abnormalities and determined there to be a stronger correlation with BCTQ-SSS items that were identified to be "primary" symptoms (i.e. numbness, tingling and nocturnal symptoms).

The results from the BCTQ-FN may be reflective of the differences between the general and restricted groups. The restricted group was comprised of individuals who must have positive EDX findings whereas the general group was comprised of subjects with a symptomatic presentation of CTS only. The greater correlation of the BCTQ-FN scores to the general group may indicate a higher overall perceived functional disability in this group.

The DASH has not been previously examined in the context of diagnostic validity. This is an instrument typically used as an outcomes measure for upper extremity complaints originating from a variety of different sources. The results from this study do not support its use as a diagnostic predictor of CTS (refer to Table 11, p. 29).

Through the transformation of the continuous variables into dichotomous variables there is a loss of the discriminative properties of the test. The dichotomization of continuous measures was necessary in order to allow for the calculation of sensitivity, specificity and likelihood ratios with a 2x 2 contingency table. This transformation of continuous measures was primarily done by calculating ROC curves and determining a cut-point that maximized both sensitivity and specificity. Had a different cut-point been selected then the calculated validity values would have differed. Riddle and Statford [72] suggest the use of a "validity index" whereby likelihood ratios are determined at several different levels of a continuous measure, thus preserving the discriminative properties of the measure.

On the basis of an estimated prevalence or pre-test probability for CTS of 66.2% ("general" classification) or 51.4% ("restricted" classification) in this sample, LR+ values exceeding 2 and LR- values less than 0.5 (Table 10) would result in post-test probability changes of at least 13.5% for the "general" group and 16.5% for the "restricted" group (as per formulae Appendix C). Tests with the highest +LRs provide

the most information in the event of a positive test, and tests with the lowest –LRs provide the most information in the event of a negative test.[48] Table 13 presents the calculated post-test odds for the two tests with the highest +LRs and for the two tests with the lowest –LRs for both the general and restricted diagnostic classification groups. Clinical tests that result in a large shift in post-test probability are more useful in clinical decision making: they provide clinicians with more information on the likelihood of a particular diagnosis based on either a positive or negative test result. The absence of hand numbness (SSS-6), for example, is, among the items examined in this study, the best predictor of ruling out a diagnosis of CTS using both the general and restricted classifications as it produces the largest shift in post-test probability based on very low negative LR values.

Table 13 : Calculated Post Test Probabilities

Test	**LR Value**	Post-Test Probability	Change in probability
General Classification:			
Flick Sign	7.65	93.76%	27.56%
Thenar Atrophy	6.25	92.45%	26.25%
SSS- 6	0.12	19.0%	47.20%
BCTQ-FN	0.37	42.13%	24.07%
Restricted Classification:			
Flick Sign	6.63	89.13%	37.73%
Thenar Atrophy	4.74	83.41%	32.01%
SSS-6	0.11	10.39%	41.01%
BCTQ-SSS	0.38	28.71%	22.69%

5.4 **Integration and Interpretation of Multiple Clinical Tests**

The interpretation of a single diagnostic test item is simple to do through the use of LRs and calculation of post-test probability of a condition given a known or estimated pre-test probability [48](Appendix C). When there are multiple clinical tests a serial multiplication of LRs can be done whereby the post-test probability calculated from the first test's LR becomes the pre-test probability for the second test, and so on. [48, 53] This method, however, assumes that the tests are conditionally independent. If two or more of

the tests are not conditionally independent then the estimate diagnostic accuracy may be inaccurate. [73]

Holleman & Simel [73] describe four strategies for integration of multiple test results: 1) using all items found to be useful by bivariate analysis, 2) using the single best LR, 3) using only items identified by a logistic regression analysis, and 4) using only items chosen by logistic regression analysis and adjusting the LRs to conform to the independence assumption. The use of the first method can result in decreased diagnostic accuracy and potential violation of the independence assumption. The remaining three strategies are proposed by Holleman & Simel as the best methods to integrate multiple clinical tests. Table 12 presents the results from this study using these three strategies. The single test item (i.e. SSS item 6= presence of numbness in hands) yielded the best diagnostic accuracy of the three methods (DOR=12.5 and 12 for general and restricted classification respectively).

The integration of the results of multiple clinical tests, which includes both physical examination and items clinical history (such as symptom presentation), is a vital component of an evidence-based clinical reasoning process. The results from this study suggest that the use of every relevant historical item and performing every possible physical examination maneuver may not only cause inefficiency, but may also create inaccurate assessments (and diagnoses).[73] The combination of multiple test results can, however, be complicated and fraught with potential bias, particularly from the violation of any required assumptions of independence. The use of logistic regression can be a way to overcome this. An underlying assumption of logistic regression; however, is that the independent variables are mutually exclusive or independent[74]. It is also proposed that a sample size of at least 50 subjects per independent variable might be required for accurate hypothesis testing[75].

Only two items were identified by the logistic regression analysis as significant predictors of a CTS diagnosis for each of the general and restricted classification schemes. Symptom provocation with the ULNT1 and the presence of numbness as a symptom were identified as significant predictors of a “general” CTS classification. The adjusted TIC validity (DOR=12.5) was identical to the validity of the SSS item 6 (i.e. single best test item). The combined validity of both items, TIC 2 (DOR=9.63), was

higher than the TIC 1 (DOR=1.73). Thus the most valid item, related to a CTS general classification, is SSS Item 6 on its own; however, symptom provocation with the ULNT 1 combined with the presence of hand numbness as a symptom (SSS Item 6) appear to have promise with respect to CTS diagnosis. With a restricted CTS classification, SSS item 6 demonstrated the best validity (DOR=12). The validity of both items identified (i.e. TIC 2) was identical to the use of the "adjusted" TIC (DOR=6). Had more variables been identified by the logistic regression analysis then an "adjusted" model would have likely resulted in a different level of validity than the use of unadjusted variables simply identified by the multivariate analysis.

Wainner and colleagues[56] developed a clinical prediction rule for the diagnosis of carpal tunnel syndrome that was substantially different from the TICs determined in this study. The results of their study showed that if five items (the BCTQ-SSS score >1.9, a wrist ration index of >.67, a patient report of shaking the hand for symptom relief (i.e. flick sign), diminished sensation of the thumb pad, and age>45) were present the +LR was 18.3 (95%CI: 1.0, 328.3). Some of the differences may be accounted for in study design, and the application and interpretation of clinical tests and self-report items. The reference standard used in Wainner et al's [56] study was EDX results only in contrast to this study where both symptom presentation and EDX results were considered. Their study did not account for bilateral symptom presentation in that only the "more symptomatic limb" was considered the involved side. The "more symptomatic limb" as interpreted by the subjects may not reflect the relative severity of the condition as decreased severity of symptoms, such as paresthesia, tends to occur in more severe cases of CTS[76]. Their method of sensation testing had not been previously reported thus comparison is difficult. Their interpretation of ULNT positivity included any one of three conditions (i.e. symptom reproduction, difference in ROM of > 10 degrees, increased or decreased symptoms with neck side flexion away or towards the tested limb). Their study did not find the ULNT1 to be diagnostically valid; however in this study ULNT1 symptom reproduction was determined to be diagnostically valid for both the general and restricted classification groupings and was an identified item of the general TIC. This study examined the three condition of ULNT1 positivity separately and thus this may account for the difference. This study examined the total average

score of the BCTQ-SSS and each individual item whereas Wainner et al[56] examined only the BCTQ-SSS total score; examination of the individual items of the BCTQ-SSS in this study lead to determining the relative predictive value of "hand numbness" for the diagnosis of carpal tunnel syndrome.

The development of a clinical prediction rule that can be readily and accurately applied in a variety of clinical settings is a multi-step process [57]. The first step is the identification of items that comprise the TIC[77]. The next step entails the validation of the identified TIC in a different group of patients by a different group of researchers in a variety of different clinical settings [77]. This study and Wainner et al's[56] efforts are but the first step in this process.

5.5 **Study Strengths and Limitations**

The strength of this study lies primarily in the design. A prospective blinded comparison of the clinical tests to a reference standard applied to a consecutive series of patients from a relevant clinical population minimizes many of the potential sources of bias that can arise from a study of this nature. This study examines both the reliability and the validity of the clinical tests which is essential to determining a complete picture of the potential utility of a clinical test. The inclusion of both traditional (e.g. Phalen's, Tinel's) and non-traditional clinical tests (e.g. Flick sign, ULNT1) and self-report measures (BCTQ and DASH) has not previously been done within the context of CTS diagnosis (with the exception of Wainner et al's study[56]) The use of two CTS classification schemes (i.e. general and restricted) allows for a greater depth of interpretation of validity and attempts to overcome the limitations of the use of EDX tests on their own as a "gold standard."

The greatest limitation of this study is the restricted sample size. The wide confidence intervals for most measures attest to the imprecision of the reliability and validity estimates. An a priori sample size calculation could have been done do determine the optimum sample size; however, the number of subjects was primarily based on the restricted time line of the study. The reliability of some clinical tests was quite poor. The inter-rater reliability of ULNT1 ROM measure in particular was low in comparison to previous research; this may have ultimately affected the validity of this

test. The ULNT ROM validity was determined on both the basis of an ROC curve cut-point and the use of a 10 degree difference from side to side. The latter method is likely not an appropriate method of determining positivity in a clinical group that has such a high prevalence of bilateral symptom presentation. The dichotomization of continuous variables in this study results in a loss of the discriminative properties of the measure which perhaps could have been preserved had a "validity index"[72] been used. The use of logistic regression to determine the TICs was likely not the ideal method of analysis in this study given the small sample size and the probable violation of the independence assumption.

5.6 **Future Directions**

A study using a larger sample size is required to increase the precision of the diagnostic validity point estimates obtained in the study and to validate the items and properties of the TICs produced. The use of a validated TIC could serve as the basis for determination of a homogenous group for clinical trials. A study examining the effectiveness of neuromobilization techniques in patients with CTS demonstrated by the combination of numbness and symptom provocation with ULNT 1 may, for example, be warranted. Examination of means to optimize the reliability of some of the clinical tests (e.g. ULNT 1 ROM) may result in greater validity and overall diagnostic utility. Also the determination of normative ROM values for the ULNT would lead to improved determination of what an "abnormal" test might be, particularly when faced with a bilateral symptom presentation. A trial assessing patient outcomes and cost outlays to determine whether patients are benefited by undergoing the clinical tests and self-report measures vs. traditional EDX testing may also be warranted.

6. **CONCLUSIONS**

Most individual items of the clinical examination in this study were found to have at least a fair level of reliability. Those with poor levels of reliability, such as the ULNT 1 ROM, had many potential sources of variability due to instrument, tester and subject influences. Optimization of reliability may result in improved diagnostic accuracy of some tests.

The clinical test items that had acceptable diagnostic validity with a "general" CTS classification were: Phalen's, carpal compression test, grip strength, thumb abduction strength, flick sign, thenar atrophy, and ULNT1 symptom reproduction. The clinical test items that had acceptable diagnostic validity with a "restricted" CTS classification were: flick sign, thenar atrophy, and ULNT symptom reproduction. For most clinical tests the validity values calculated for the general CTS diagnostic classification were better than those determined for the restricted classification.

The findings of this study do not support the use of the DASH as a diagnostic tool for CTS. The functional scale of the BCTQ was the only questionnaire that had acceptable validity with a general CTS classification. Only the symptom severity scale of the BCTQ had acceptable diagnostic validity with a restricted CTS classification. The individual items of the BCTQ-SSS that had acceptable diagnostic validity were item 1 ("wrist and hand pain at night"- general classification only), item 6 ("hand numbness"- general and restricted classification), and item 11 ("difficulty with grasping and use of small objects"- restricted classification only). Hand numbness (SSS item 6) had the best validity of any of the individual clinical tests, questionnaires or items from the BCTQ-SSS for both restricted and general classification schemes.

The TIC derived through logistic regression for the general classification group was SSS item 6 and symptom provocation with ULNT1 and the TIC determined for the restricted classification group was SSS item 6 and BCTQ-SSS average score. The TICs for both classification groupings did not yield improved diagnostic validity beyond that found with the single best test item (SSS item 6).

Based on the results of this study the best diagnostic "test" for CTS is not a clinical test, per se, but a simple item of clinical history. Item 6 of the BCTQ-SSS ("hand

numbness") demonstrated the best diagnostic validity when compared to all other individual test items and the derived TIC's for both the general and restricted classification schemes. This item had a low –LR value, thus the diagnostic power of this test lies in a negative test result or in the "absence" of hand numbness.

The process of diagnosis is an essential part of successful clinical management of a patient. The results of individual tests can be used to determine which conditions can be ruled in or out, ultimately leading to a decision as to which interventions will provide optimum patient outcomes. Clinicians are presented with a plethora of items of history and clinical tests are whenever a patient is examined. An evidence-based clinical reasoning approach allows a clinician to sift through and priorize between these multiple items to determine the likelihood of a diagnosis. Even though this study showed one item (i.e. "hand numbness") to be diagnostically superior to all other items examined for the diagnosis of CTS, the use of only one test item in a clinical scenario is not practical. Instead, the findings of this study can be used to assist clinicians determine which items of the examination have more or less value for the diagnosis of CTS.

7. **REFERENCES**

[1] D'Arcy CA, McGee S. Does this patient have carpal tunnel syndrome? Journal of American Medical Association. 2000;283:3110-3117.

[2] Phalen GS. The carpal tunnel syndrome: seventeen years' experience in diagnosis and treatment of six hundred fifty-four hands. J Bone Joint Surg Am. 1966;48:211-228.

[3] Stevens JC, Smith BE, Weaver AL et al. Symptoms of 100 patients with electromyographically verified carpal tunnel syndrome. Muscle and Nerve. 1999;22:1448-1456.

[4] Ashworth NL. Carpal tunnel syndrome. E medicine. 2002. www.emedicine.com/pmr/topic21.htm.

[5] Atroshi I, Gummesson C, Johnsson R, et al. Prevalence of carpal tunnel syndrome in a general population. JAMA. 1999; 282(2): 153-8.

[6] de Krom MC, Kester AD, Knipschild PG, Spaans F. Risk factors for carpal tunnel syndrome. Epidemiology. 1990 Dec; 132(6): 1102-10.

[7] Papanicolaou GD, McCabe SJ, Firrell J. The prevalence and characteristics of nerve compression symptoms on the general population. J Hand Surg (Am). 2001 May; 26(3): 460-6.

[8] Mondelli M, Giannini F, Giacchi M. Carpal tunnel syndrome incidence in a general population. Neurology. 2002 Jan; 58(2): 289-94.

[9] Nordstrom DL, DeStefano F, Vierkant RA, Layde PM. Incidence of diagnosed carpal tunnel syndrome in a general population. Epidemiology 1998 May; 9(3): 342-5.

[10] Garland FC, Garland CF, Doyle EJ Jr et al. Carpal tunnel syndrome and occupation in US Navy enlisted personnel. Arch Environ Health 1996 Sep-Oct; 51(5): 395-407.

[11] Falkiner S, Myers S. When exactly can carpal tunnel syndrome be considered work related? ANZ J Surg. 2002 Mar; 72(3): 204-9.

[12] Goga IE. Carpal tunnel syndrome in black South Africans. J Hand Surg (Br) 1990 Feb; 15 (1): 96-99.

[13] Verdugo RJ, Salinas RS, Castillo J, Cea JG. Surgical versus non-surgical treatment for carpal tunnel syndrome The Cochrane Database of Systematic Reviews 2003, Issue 3. Art. No.: CD001552. DOI: 10.1002/14651858.CD001552.

[14] Katz JN, Carpal tunnel syndrome: a practical review. American Family Physician. 1994; 49(6): 1371-1379.

[15] Palmer AK, Toivonen DA. Complications of endoscopic and open carpal tunnel release. Southern Medical Journal. 1999; 87(3); 416-418.

[16]Kluge W, Simpson RG, Nicol AC. Late complications after open carpal tunnel decompression. Journal of Hand Surgery (Am). 1996; 23(4): 697-710.

[17] Marshall S, Tardif G, Ashworth N. Local corticosteroid injection for carpal tunnel syndrome (Cochrane Review). In: *The Cochrane Library*, Issue 1, 2004. Chichester, UK: John Wiley & Sons, Ltd.

[18] O'Connor D, Marshall S, Massy-Westropp N. Non-surgical treatment (other than steroid injection) for carpal tunnel syndrome (Cochrane Review). In: *The Cochrane Library*, Issue 1, 2004. Chichester, UK: John Wiley & Sons, Ltd.

[19] Tal-Akabi A, Rushton A. An investigation to compare the effectiveness of carpal bone mobilization and neurodynamic mobilization as methods for treatment for carpal tunnel syndrome. Manual Therapy. 2000; 5(4): 214-222.

[20] Akalin E et al. Treatment of carpal tunnel syndrome with nerve and tendon gliding exercises. Am J Phys Med Rehab. 2002; 81(2): 108-113.

[21] Padua L, Padua R, Nazzaro M, Tonali P. Incidence of bilateral symptoms in carpal tunnel syndrome. J Hand Surg (Br). 1998 Oct; 23(5): 603-6.

[22] Bodofsky EB, Greenberg WM, Wu KD. Median nerve compression at the wrist: is it ever unilateral? Electromyogr Clin Neurophysiol. 2001 Dec; 41(8): 451-6.

[23] Upton AR, McComas AJ. The double crush in nerve entrapment syndromes. Lancet. 1973 Aug; 2 (7825): 359-62.

[24] Pierre-Jerome C, Bekkelund SI. Magnetic resonance assessment of the double-crush phenomenon in patients with carpal tunnel syndrome: a bilateral quantitative study. Scand J Palst Reconstr Surg Hand Surg. 2003; 37(1): 46-53.

[25] Slater H, Butler D, Shacklock MO. The dynamic central nervous system; examination and assessment using tension tests. In: Boyling J. Palastanga N (Eds) Grieve's Modern Manual Therapy, 2nd Edn. Churchill Livingstone: Edinburgh; pp 587-606.

[26] Mackinnon SE, Dallon A, Phillipe JH. Chronic human nerve compression- a histological assessment. Neuropathology and Applied Neurobiology. 1986; 12.

[27]Greening J, Smart S, Leary R et al. Reduced movement of median nerve in carpal tunnel during wrist flexion in patients with non-specific arm pain. The Lancet. 1999 July; 354(9174): 217-8.

[28] Nakamichi K, Tachibana S. Restricted motion of the median nerve in carpal tunnel syndrome. Journal of Hand Surgery (Br). 1995; 20B(4): 460-4.

[29] Valls-Sole J, Ramiro A, Nunez M. Limited longitudinal sliding of the median nerve in patients with carpal tunnel syndrome. Muscle and Nerve. 1995; 18: 761-7.

[30] Butler DS. The Sensitive Nervous System. Noigroup Publications, Australia, 2000.

[31] Elvey RL. Brachial plexus tension tests and the pathoanatomical origin of arm pain. In: Idczak R (ed) Aspects of Manipulative Therapy, Manual Physiotherapists Association of Australia, Melbourne. 1979.

[32] Shacklock MO. Positive upper limb tension test in a case of surgically proven neuropathy: analysis and validity. Manual Therapy. 1996 June; 1(3): 154-161.

[33] Rempel D, Evanoff B, Amadio PC et al. Consensus criteria for the classification of carpal tunnel syndrome in epidemiological studies. American Journal of Public Health. 1998 Oct; 88(10): 1447-1451.

[34] Ferry S, Silman AJ, Pritchard T, Keenan J, Croft P. The association between different patterns of hand symptoms and objective evidence of median nerve compression: a community-based survey**.** Arthritis Rheum. 1998 Apr; 41(4): 720-4.

[35] Rosthein JM, Campbell SK, Echternach JL, et al. Standards for tests and measurement in physical therapy practice. Physical Therapy. 1991; 71: 589-622.

[36] Domholdt E. Physical Therapy Research Principles and Applications. Toronto: W.B. Saunders Company, 1993: 168-170.

[37] MacDermid JC, Kramer JF, McFarlane RM, Roth JH. Inter-rater agreement and accuracy of clinical tests used in diagnosis of carpal tunnel syndrome. Work A Journal of Prevention & Rehabilitation. 1997; 8: 37-44.

[38] Sackett DL, Haynes RB, Tugwell P. Clinical Epidemiology. A Basic Science for Clinical Medicine, Little Brown and Company, Toronto, 1985.

[39] Buch-Jaeger N, Foucher, G. Correlation of clinical signs with nerve conduction tests in the diagnosis of carpal tunnel syndrome. Journal of Hand Surgery (British). 1994; 19(6): 720-724.

[40] Kulhman KA, Hennessey WJ. Sensitivity and specificity of carpal tunnel syndrome signs. American Journal of Physical Medicine and Rehabilitation. 1997; 76(6): 451-457.

[41] Szabo, RM, Slater RR, Farver TB et al. The value of diagnostic testing in carpal tunnel syndrome. The Journal of Hand Surgery. 1999; 24(A): 704-714.

[42] Marx, RG, Bombardier C, Wright JG. What do we know about the reliability and validity of physical examination tests results used to examine the upper extremity? The Journal of Hand Surgery. 1999; 24(A): 185-193.

[43] Coppieters M, Stappaerts K, Janssens K. Reliability of detecting "onset of pain" and "submaximal pain" during neural provocation testing of the upper quadrant. Physiotherapy Research International. 2002; 7(3): 146-156.

[44] Solway S, Beaton DE, McConnell S, Bombardier C. The DASH outcome measure user's manual, Second edition. Toronto: Institute for Work & Health, 2002.

[45] Beaton DE, Katz JN, Fossel AH et al. Measuring the whole or the parts? Validity, reliability and responsiveness of the Disabilities of the Arm, Shoulder and Hand outcome measure in different regions of the upper extremity. 2001; 14(2): 128-46.

[46] Levine DW, Simmons BP, Koris MJ et al. A self-administered questionnaire for the assessment of severity of symptoms and functional status in carpal tunnel syndrome. The Journal of Bone and Joint Surgery. 1993; 75(11): 1585-92.

[47] Fritz JM, Wainner RS. Examining Diagnostic Tests: An Evidence Based Perspective. Physical Therapy 2001; 81(9): 1546-1564.

[48] Sackett DL, Straus SE, Richardson WS et al. Evidence-Based Medicine- How to Practice and Teach EBM. London: Churchill Livingstone, 2000: 67-80.

[49] Hagen MD. Test Characteristics: How good is that test? Medical Decision Making.1995; 22:213-233.

[50] Dujardin B. Van den Ende J. Van Gompel A. et al. Likelihood ratios: a real improvement for clinical decision making? European Journal of Epidemiology. 1994; 10: 29-36.

[51] Lurie JD, Sox HC. Principles of medical decision making: spine update. Spine. 1999; 24: 493-498.

[52] Simel DL, Samsa GP, Matchar DB. Likelihood ratios with confidence: sample size estimation for diagnostic test results. Journal of clinical epidemiology. 1991; 44: 763-770.

[53] Davidson M. The interpretation of diagnostic tests: a primer for physiotherapists. Australian Journal of Physiotherapy. 2002; 48: 227-233.

[54] Lijmer JG, Mol BW, Heisterkamp S, et al. Empirical evidence of design-related bias in studies of diagnostic tests. JAMA. 1999; 282:1061–1066.

[55] Glas AS. Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PM. The diagnostic odds ratio as a single indicator of test performance. Journal of Clinical Epidemiology. 2003; 56(11): 1129-35.

[56] Wainner RS, Fritz JM, Irrgang JJ, et al. Development of a clinical prediction rule for the diagnosis of carpal tunnel syndrome. Archives of Physical Medicine and Rehabilitation. 2005; 86: 609-618.

[57] Cleland J. Orthopaedic Clinical Examination: An Evidence-Based Approach for Physical Therapists. New Jersey, Icon Learning Systems, 2005: 16.

[58] Wainner RS, Fritz JM, Irrgang JJ et al. Reliability and Diagnostic Accuracy of the Clinical Examination and Patient Self-Report Measures for Cervical Radiculopathy. Spine. 2003; 28 (1): 52-62.

[59] Stevens CJ. AAEM Minimograph #36: The electrodiagnosis of carpal tunnel syndrome. Muscle & Nerve. 1997: 1477-1486.

[60] Domholdt E. Physical Therapy Research Principles and Applications. Toronto: W.B. Saunders Company, 1993: 287-289.

[61] Hazard Munroe B. Statistical Methods in Health Care Research. 4th ed.: Philadelphia: Lippincott, 2001: 236.

[62] Hanley JA, McNeil BJ. The meaning and use of the area under a received operating characteristic (ROC) curve. Radiology. 1982; 143: 29-28.

[63] Hosmer DW, Lemeshow S. Applied Logistic Regression. New York: John Wiley, 1989: 142-4.

[64] Kenney WL, Humphrey RH, Bryant CX eds. ACSM's Guidelines for Exercise Testing and Prescription, Baltimore: Williams & Wilkins, 1995: 58-9.

[65] Irwig L, Tosteson ANA, Gatsonis C et al. Guidelines for meta-analyses evaluating diagnostic tests. Annals of Internal Medicine. 1994; 120: 667-676.

[66] Panzer RJ. Suchman AM, Griner PF. Workup bias in prediction research. Medical Decision Making. 1987; 7: 115-119.

[67] Reid MC, Lachs MS, Feinstein AR. Use of methodological standards in diagnostic test research: getting better but still not good. JAMA. 1995; 274: 645-651.

[68] Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature III: how to use an article about a diagnostic test and are the results of the study valid? JAMA. 1994; 271: 389-391.

[69] Domholdt E. Physical Therapy Research Principles and Applications. Toronto: W.B. Saunders Company, 1993: 162-171.

[70] Greenhalgh T. How to read a paper: papers that report diagnostic or screening tests. BMJ. 1997; 315: 540-543.

[71] You H, Simmons Z, Freivalds A et al. Relationships between clinical symptom severity scales and nerve conduction measures in carpal tunnel syndrome. Muscle & Nerve. 1999; 22: 497-501.

[72] Riddle DL & Stratford PW. Interpreting validity indexes for diagnostic tests: An illustration using the Berg Balance Test. Physical Therapy. 1999;79 (10): 939-948.

[73] Holleman DR, Simel DL. Quantitative assessment from the clinical examination: how should clinicans integrate numerous results? Journal of General Internal Medicine. 1997; 12: 165-171.

[74] Kleinbaum DG, Klein M. Logistic Regression A Self-Learning Text 2nd ed. Springer. New York 2002.

[75] Grimm LG & Yarnold, PR eds. Reading and Understanding Multivariate Statistics. American Psychological Association. Washington D.C. 1995.

[76] MacDermid JC, Doherty T. Clinical Electrodiagnostic Testing of Carpal Tunnel Syndrome: A Narrative Review. Journal of Orthopaedic & Sports Physical Therapy. 2004; 34: 565-588.

[77] McGinn T Guyatt G, Wyer P et al. Users' guide to the medical literature XXII: how to use articles about clinical decision rules. JAMA. 2000; 284: 79-84.

[78] Kendall FP, McCreary EK, Provance PG. Muscle Testing and Function 4th ed. Williams & Wilkins. Baltimore. 1993.

8. **APPENDICES**

for a Masters Thesis

Reliability and Diagnostic Validity of Clinical Examination and Patient Self-Report Measures in Carpal Tunnel Syndrome

Brenna Bath, (MSc Student)

Table of Appendices

APPENDIX A- DISABILITY OF ARM SHOULDER AND HAND QUESTIONNAIRE (DASH) [43]

Description of Tool:

The DASH is a 30-item self-report condition-specific disability measure. The questionnaire includes 2 physical function items, 6 symptom items and 3 social/role function items. Each item of the DASH is scored on a 5 point scale (1 to 5).

Application/Administration of Tool:

Subjects circle the appropriate response to each question based on their condition in the last week. If the subject did not perform a specific activity in the past week, they are asked to make their best estimate of which response would be most accurate. To calculate the overall DASH score, responses to the 30-item DASH are summed. The sum is then transformed so that it falls between 0 and 100 by subtracting 30 and dividing by 1.2. If less than 3 items are blank, the mean score of the other items may be substituted for the missing scores. If four or more scores are left blank, a DASH score cannot be calculated.

Interpretation of Tool:

Lower scores reflect less disability and higher scores reflect more disability. Preliminary evidence suggests that individuals who are able to do their work despite upper extremity pain, tend to score approximately 20-30 on the DASH, in contrast to scores in the 50-80 range for those unable to work because of their arm problem.

DISABILITY OF ARM SHOULDER AND HAND QUESTIONNAIRE (DASH)

INSTRUCTIONS

- This questionnaire asks about your symptoms as well as your ability to perform certain activities.
- Please answer every question, based on your condition in the last week, by circling the appropriate number.
- If you did not have the opportunity to perform an activity in the past week, please make your best estimate on which response would be the most accurate.
- It doesn't matter which hand or arm you use to perform the activity; please answer based on your ability regardless of how you perform the task.

Please rate your ability to do the following activities in the last week by circling the number below the appropriate response.

	NO DIFFICULTY	MILD DIFFICULTY	MODERATE DIFFICULTY	SEVERE DIFFICULTY	UNABLE
1. Open a tight or new jar.	1	2	3	4	5
2. Write.	1	2	3	4	5
3. Turn a key.	1	2	3	4	5
4. Prepare a meal.	1	2	3	4	5
5. Push open a heavy door.	1	2	3	4	5
6. Place an object on a shelf above your head.	1	2	3	4	5
7. Do heavy household chores (e.g., wash walls, wash floors)	1	2	3	4	5
8. Garden or do yard work.	1	2	3	4	5
9. Make a bed.	1	2	3	4	5
10. Carry a shopping bag or briefcase.	1	2	3	4	5
11. Carry a heavy object (over 10 lbs).	1	2	3	4	5
12. Change a lightbulb overhead.	1	2	3	4	5
13. Wash or blow dry your hair.	1	2	3	4	5
14. Wash your back.	1	2	3	4	5
15. Put on a pullover sweater.	1	2	3	4	5
16. Use a knife to cut food.	1	2	3	4	5
17. Recreational activities which require little effort (e.g., cardplaying, knitting, etc.).	1	2	3	4	5
18. Recreational activities in which you take some force or impact through your arm, shoulder or hand (e.g., golf, hammering, tennis, etc.).	1	2	3	4	5

19. Recreational activities in which you move your arm freely (e.g., playing frisbee, badminton, etc.).	1	2	3	4	5
20. Manage transportation needs (getting from one place to another).	1	2	3	4	5
21. Sexual activities.	1	2	3	4	5

	NOT AT ALL	SLIGHTLY	MODERATELY	QUITE A BIT	EXTREMELY
22. During the past week, to what extent has your arm, shoulder or hand problem interfered with your normal social activities with family, friends, neighbours or groups?	1	2	3	4	5

	NOT LIMITED AT ALL	SLIGHTLY LIMITED	MODERATELY LIMITED	SEVERELY LIMITED	UNABLE
23. During the past week, were you limited in your work or other regular daily activities as a result of your arm, shoulder or hand problem? (circle number)	1	2	3	4	5

Please rate the severity of the following symptoms in the last week. (circle number)

	NONE	MILD	MODERATE	SEVERE	EXTREME
24. Arm, shoulder or hand pain.	1	2	3	4	5
25. Arm, shoulder or hand pain when you performed any specific activity.	1	2	3	4	5
26. Tingling (pins and needles) in your arm, shoulder or hand.	1	2	3	4	5
27. Weakness in your arm, shoulder or hand.	1	2	3	4	5
28. Stiffness in your arm, shoulder or hand.	1	2	3	4	5

	NO DIFFICULTY	MILD DIFFICULTY	MODERATE DIFFICUTLY	SEVERE DIFFICULTY	SO MUCH DIFFICUTLY THAT I CAN'T SLEEP
29. During the past week, how much difficulty have you had sleeping because of the pain in your arm, shoulder or hand?(circle number)	1	2	3	4	5

	STRONGLY DISAGREE	DISAGREE	NEITHER AGREE NOR DISAGREE	AGREE	STRONGLY AGREE
30. I feel less capable, less confident or less useful because of my arm, shoulder or hand problem. (circle number)	1	2	3	4	5

APPENDIX B- BRINGHAM CARPAL TUNNEL QUESTIONNAIRE (BCTQ) [44]

Description of Tool:

The BCTQ is a self-administered questionnaire for the assessment of severity of symptoms and functional status in patients who have CTS. The questionnaire includes a symptom severity scale with eleven questions covering six "critical" domains (pain, parasthesia, numbness, weakness, nocturnal symptoms and over-all functional status). It also includes eight items pertaining to functional status (i.e. functional status scale).

Application/Administration of Tool:

For both the symptom severity scale and the functional status scale, the subjects are asked to base their answers on a typical twenty-four hour period during the past two weeks. The symptom severity scale items are rated from 1 point (mildest) to 5 points (most severe). The answers for the functional status scale are rated from 1 point (no difficulty with the activity) to 5 points (cannot perform the activity at all).

Interpretation of Tool:

The overall symptom severity score is calculated as the mean of the scores for the eleven items. The over-all score for functional status is calculated as the mean of all eight items. Items that are left unanswered are not included in the calculation of the overall score. A lower score represents low reported symptom severity and higher functional status.

CARPAL TUNNEL QUESTIONNAIRE

Part 1: Symptom Severity Scale:

The following questions refer to your symptoms for a typical twenty-four hour period during the past two weeks (circle one answer to each question).

- How severe is the hand or wrist pain that you have a night?
  1 I do not have hand or wrist pain at night
  2 Mild pain
  3 Moderate pain
  4 Severe pain
  5 Very severe pain

- How often did hand or wrist pain wake you up during a typical night in the past two weeks?
  1 Never
  2 Once
  3 Two or three times
  4 Four or five times
  5 More than five times

- Do you typically have pain in your hand or wrist during the daytime?
  1 I never have pain during the day.
  2 I have mild pain during the day.
  3 I have moderate pain during the day.
  4 I have severe pain during the day.
  5 I have very severe pain during the day.

- How often do you have hand or wrist pain during the daytime?
  1 I never have pain during the day.
  2 Once or twice a day
  3 Three to five times a day
  4 More than five times a day
  5 The pain is constant

- How long, on average, does an episode of pain last during the daytime?
  1 I never get pain during the day.
  2 Less than 10 minutes
  3 10 to 60 minutes
  4 Greater than 60 minutes
  5 The pain is constant throughout the day.

- Do you have numbness (loss of sensation) in your hand?
  1 No
  2 I have mild numbness
  3 I have moderate numbness
  4 I have severe numbness
  5 I have very severe numbness

- Do you have weakness in your hand or wrist?
  1 No weakness
  2 I have mild weakness
  3 I have moderate weakness
  4 I have severe weakness
  5 I have very severe weakness

- Do you have tingling sensations in your hand?
  1 No tingling
  2 I have mild tingling
  3 I have moderate tingling
  4 I have severe tingling
  5 I have very severe tingling

- How severe is numbness (loss of sensation) or tingling at night?
  1 I have no numbness or tingling at night
  2 Mild
  3 Moderate
  4 Severe
  5 Very severe

- How often did hand numbness or tingling wake you up during a typical night during the past two weeks?
  1 Never
  2 Once
  3 Two or three times
  4 Four or five times
  5 More than five times

- Do you have difficulty with the grasping and use of small objects such as key or pens?
  1 No difficutly
  2 Mild difficulty
  3 Moderate difficulty
  4 Severe difficulty
  5 Very severe difficulty

Part 2: Functional Status Scale

On a typical day during the past two weeks have hand and wrist symptoms caused you to have any difficulty doing the activities listed below? Please circle one number that best describes your ability to do the activity.

- Writing
  1 No difficulty
  2 Mild difficulty
  3 Moderate difficulty
  4 Severe difficulty
  5 Cannot do at all due to hand and wrist symptoms

- Ironing of clothes
  1 No difficulty
  2 Mild difficulty
  3 Moderate difficulty
  4 Severe difficulty
  5 Cannot do at all due to hand and wrist symptoms

- Holding book while reading
  1 No difficulty
  2 Mild difficulty
  3 Moderate difficulty
  4 Severe difficulty
  5 Cannot do at all due to hand and wrist symptoms

- Gripping of a telephone handle
  1 No difficulty
  2 Mild difficulty
  3 Moderate difficulty
  4 Severe difficulty
  5 Cannot do at all due to hand and wrist symptoms

- Opening of jars
  1 No difficulty
  2 Mild difficulty
  3 Moderate difficulty
  4 Severe difficulty
  5 Cannot do at all due to hand and wrist symptoms

- Household chores
  1 No difficulty
  2 Mild difficulty
  3 Moderate difficulty
  4 Severe difficulty
  5 Cannot do at all due to hand and wrist symptoms

- Carrying of grocery bags
  1 No difficulty
  2 Mild difficulty
  3 Moderate difficulty
  4 Severe difficulty
  5 Cannot do at all due to hand and wrist symptoms

- Bathing and dressing
  1 No difficulty
  2 Mild difficulty
  3 Moderate difficulty
  4 Severe difficulty
  5 Cannot do at all due to hand and wrist symptoms

APPENDIX C–FORMULAE FOR ANALYSIS

Contingency Table Created by comparing the Results of the Diagnostic Test and the Reference Standard:

	Reference Standard Positive	Reference Standard Negative
Diagnostic test positive	True positive results (a)	False positive results (b)
Diagnostic test negative	False negative results (c)	True negative results (d)

- Sensitivity (%) = a/(a + c) x 100
  - Given that the individual has the condition, the probability that the test will be positive

- Specificity (%) = d/(d + b) x 100
  - Given that the individual does not have the condition, the probability that the test will be negative

- Positive likelihood ratio = sensitivity/(1-specificity)
  - Given a positive test result, the increase in odds favoring the condition

- Negative likelihood ratio = (1-sensitivity/specificity)
  - Given a negative test result, the decrease in odds favoring the condition

- Diagnostic Odds Ratio = (a*d)/(b*c)
  - Odds for a positive test results in diseased persons relative to odds of positive test results in non-diseased persons

- Calculation of post-test probability [51]:
  - Estimate the pre-test probability (or use known prevalence)
  - Convert to pre-test odds by dividing probability by 1-probability
  - Multiply the pre-test odds by the positive or negative LR value
  - Convert the post-test odds to post-test probability by dividing the odds by odds + 1.

- Use of a likelihood ratio nomogram [46,51]:
  - Mark the pre-test probability
  - Mark the LR of the selected tests on the middle line
  - Connect the marks and draw a line which will provide an estimate of post-test probability.

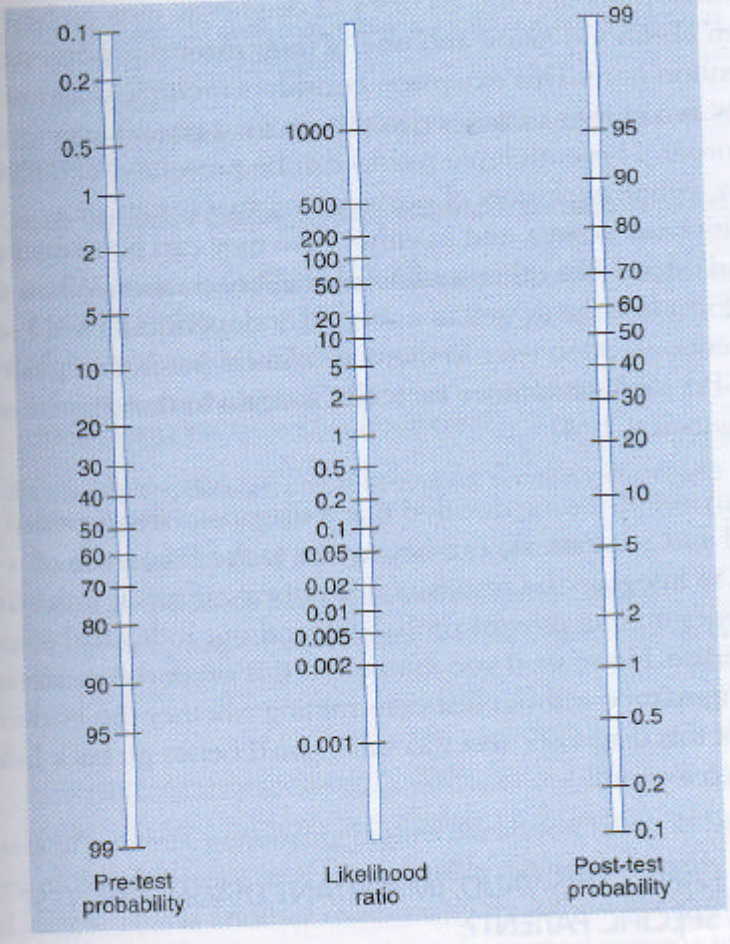


Figure 3.1 A likelihood ratio nomogram.

APPENDIX D- U of S ETHICS APPROVAL LETTER

UNIVERSITY OF SASKATCHEWAN
BEHAVIOURAL RESEARCH ETHICS BOARD
http://www.usask.ca/research/ethics.shtml

NAME: Leonard Tan (Brenna Bath) Beh 04-170
Community Health & Epidemiology

DATE: August 19, 2004

The University of Saskatchewan Behavioural Research Ethics Board has reviewed the Application for Ethics Approval for your study "Reliability and Diagnostic Accuracy of Clinical Examination and Patient Self-Report Measures in Carpal Tunnel Syndrome" (Beh04-170).

1. Your study has been APPROVED.

2. Any significant changes to your proposed method, or your consent and recruitment procedures should be reported to the Chair for Committee consideration in advance of its implementation.

3. The term of this approval is for 5 years.

4. This approval is valid for one year. A status report form must be submitted annually to the Chair of the Committee in order to extend approval. This certificate will automatically be invalidated if a status report form is not received within one month of the anniversary date. Please refer to the website for further instructions http://www.usask.ca/research/behavrsc.shtml

I wish you a successful and informative study.

Dr. Scott Bell, Acting Chair
University of Saskatchewan
Behavioural Research Ethics Board

VT/ck

Office of Research Services, University of Saskatchewan
Room 1607, 110 Gymnasium Place, Box 5000 RPO University, Saskatoon SK S7N 4J8 CANADA
Telephone: (306) 966-8576 Facsimile: (306) 966-8597
http://www.usask.ca/research

APPENDIX E- SASKATOON HEALTH REGION APPROVAL LETTER

Saskatoon Health Region

Research Services Unit
Strategic Health Information & Planning Services (SHIPS)
Joanne Franko, Manager
Suite 300 Saskatoon Square
410 22nd St E
Saskatoon, SK S7K 5T6
Phone: 306.655.3356 Fax: 306.655.3373

DATE: **September 23, 2004**

TO: **Brenna Bath, MSc Student, Community Health & Epidemiology, U of S**

FROM: **Joanne Franko**
Manager, Research Services Unit

RE: **RESEARCH PROJECT ETHICS COMMITTEE (EC)#: 2004-170**
PROJECT NAME: Reliability and Diagnostic Accuracy of Clinical Examination and Patient Self-Report Measures in Carpal Tunnel Syndrome
PROTOCOL #: N/A

Saskatoon Health Region is pleased to provide you with operational approval of the above-mentioned research project.

Please advise me when the data collection phase of the research project is completed. I would also appreciate receiving a summary of the results for this research project. As well, any publications or presentations that result from this research should include a statement acknowledging the assistance of Saskatoon Health Region.

I would like to wish you every success with your project and encourage you to contact me if I can assist you with it.

If you have any questions, please contact my office at 655-3356.

Yours truly,

Joanne Franko, M.Sc.
Manager, Research Services Unit

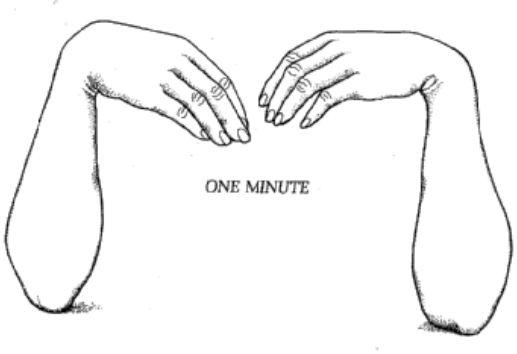
cc: Cherryl Walter, Mgr., Rehabilitation Medicine, SCH
Carol Brown, Mgr., F. I. T. For Active Living, SCH

APPENDIX F- OPERATIONAL DEFINITIONS OF CLINICAL TESTS

Provocation Tests:

Phalen sign [1]:

- **Test Description:**
  - Subject's wrists are passively flexed to 90 degrees for 60 seconds

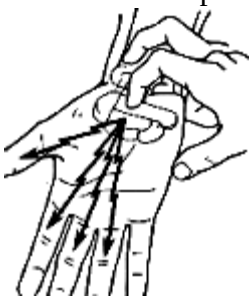


- **Definition of Abnormal Finding:**
  - Parasthesia (i.e. "pins and needles" sensation) in the distribution of the median nerve

- **Level of Measurement for Analysis:**
  - Nominal (positive or negative)

Tinel sign [1]:

- **Test Description:**
  - Clinician taps on distal palmar wrist crease over the median nerve



- **Definition of Abnormal Finding:**
  - Parasthesia in the distribution of the median nerve

- **Level of Measurement for Analysis:**
  - Nominal (positive or negative)

Carpal Compression Test [1]:

- **Test Description:**
  - Examiner presses with his/her thumb on the palmar aspect of the subject's wrist at the level of the carpal tunnel for 60 seconds

- **Definition of Abnormal Finding:**
  - Parasthesia in the distribution of the median nerve

- **Level of Measurement for Analysis:**
  - Nominal (positive or negative)

Sensation Examination:

Hypoalgesia (Sharp/dull sensibility) [1]:

- **Test Description:**
  - Examiner uses the sharp tip of an open paper clip (discarded after use to prevent infection) to apply a pressure to the subject's palmar aspect of the index finger and to the ipsilateral palmar aspect of the little finger. The subject is asked if they perceive the stimulus applied to the index finger to be less than that applied to little finger.

- **Definition of Abnormal Finding:**
  - Diminished ability to perceive painful stimuli applied to the index finger compared to the ipsilateral little finger.

- **Level of Measurement for Analysis:**
  - Nominal (positive or negative)

Monofilament Testing (pressure sensibility)[1]:

- **Test Description:**
  - Semmes-Weinstein monofilaments are applied to the pulp of the index finger. The subject is asked to close his/her eyes and report when a pressure sensation is felt.

- **Definition of Abnormal Finding:**
  - The subject's threshold of pressure sensation is greater than the 2.83 monofilament (i.e. 3.60 or above).

- **Level of Measurement for Analysis:**
  - Nominal (positive or negative)

Strength Examination:

Grip Strength [39]:

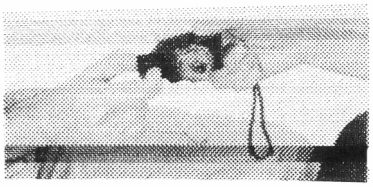
- **Test Description:**
  - A Jamar dynamometer was used to measure grip strength. The subject was instructed to grip the device and squeeze as hard as they feel comfortable with. Grip strength was measured at setting III on the dynamometer and repeated three times each side. A mean value for each side was used for analysis.

- **Definition of Abnormal Finding:**
  - Diminished strength more than 12 % less on the affected side than the contralateral side. [41] In the case of bilateral CTS, the "affected" side will be the side designated to be "worse" or more symptomatic by the subject.

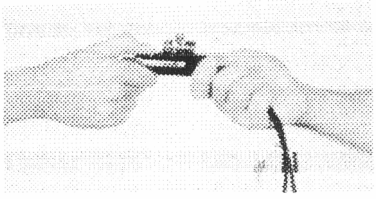
- **Level of Measurement for Analysis:**
  - Nominal (normal or diminished)

Pinch Strength [39]:

- **Test Description:**
  - A Jamar pinch dynamometer was used to measure pinch strength of three pinch types (key, tripod and tip to tip pinch). The subjects were instructed to pinch the device and squeeze as hard as they feel comfortable with. Each test was performed three times on each side.



Tip to tip pinch



Key Pinch

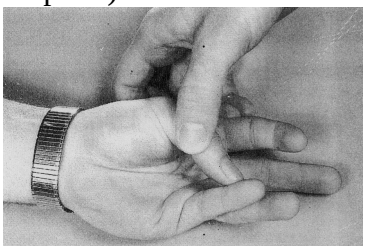
Three jaw chuck pinch

- **Definition of Abnormal Finding:**
  - Diminished strength more than 12 % less on the affected side than the contralateral side[39]. In the case of bilateral CTS, the "affected" side was the side designated to be "worse" or more symptomatic by the subject. A mean value for each side was used for analysis.

- **Level of Measurement for Analysis:**
  - Nominal (normal or diminished)

Thumb Abduction Stength [1]:

- **Test Description:**
  - The examiner resists thumb abduction (ie. movement of the thumb at right angles to the palm).



- **Definition of Abnormal Finding:**
  - Strength graded according to manual muscle testing guidelines [78].
    - Grade 5: Muscle can hold the test position against strong pressure.
    - Grade 4: Muscle can hold test position against moderate pressure.
    - Grade 3: Muscle can hold test position against gravity, but cannot hold if even slightest pressure applied.
    - Grade 2: Muscle able to move segment through a full range in a movement horizontal with respect to gravity or partial range of movement in a position against gravity.
    - Grade 1: Trace evidence of muscle contraction.
    - Grade 0: No evidence of any muscle contraction.

- **Level of Measurement for Analysis:**
  - Nominal (normal: 5 or diminished: grade 4 or lower)

Other Tests:

Flick Sign [1]:

- **Test Description:**
  - The examiner asks the subject "What do you do with your hand(s) when the symptoms are at their worst?"
- **Definition of Abnormal Finding:**
  - The subject demonstrates a flicking movement of the wrist and hand similar to that used in shaking down a thermometer.

- **Level of Measurement for Analysis:**
  - Nominal (present or absent)

Thenar Atrophy [1]:

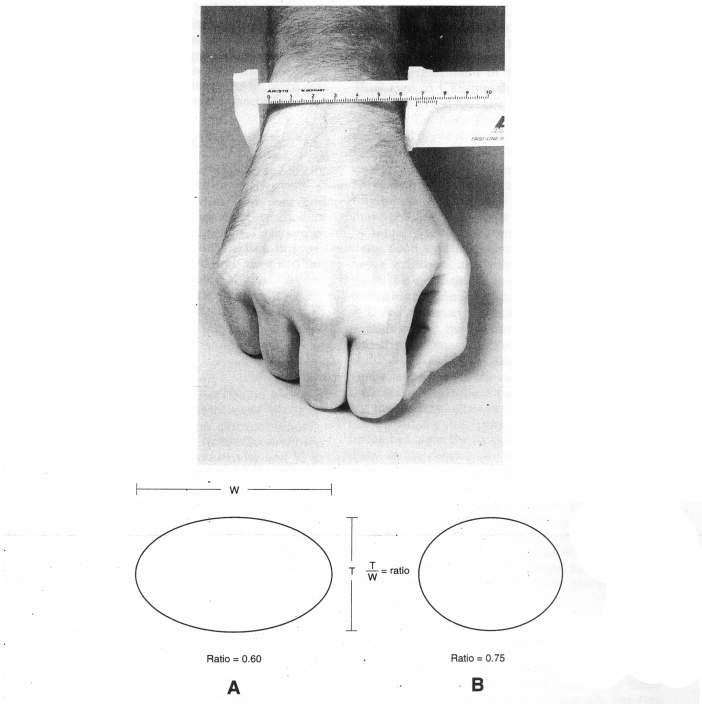
- **Test Description:**
  - The examiner visually inspects the thenar eminence from the side.

- **Definition of Abnormal Finding:**
  - A concavity of the thenar muscles when observed from the side.

- **Level of Measurement for Analysis:**
  - Nominal (present or absent)

Square wrist sign [1]:

- **Test Description:**
  - The examiner measures the anterioposterior and mediolateral dimensions of the subject's wrist with calipers at the distal wrist crease.

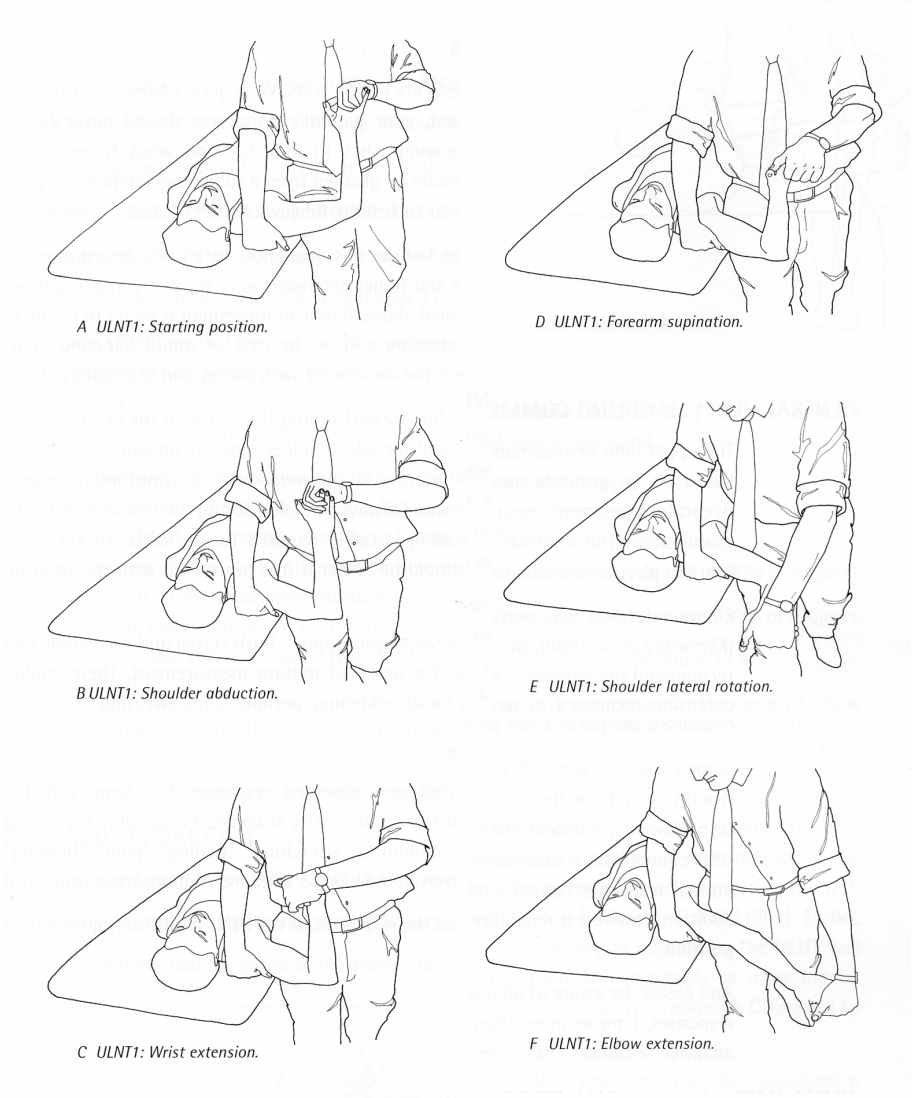


- **Definition of Abnormal Finding:**
  - The A/P dimension of the wrist divided by the M/L dimension equals a ratio of greater than 0.70.

- **Level of Measurement for Analysis:**
  - Nominal (positive or negative)

Upper Limb Neurodynamic Test 1 (ULNT 1)[30,41]:

- **Test Description:**
  - The subject starting position is in supine with both legs extended and the cervical spine in a neutral position. Any deviations from the starting position (i.e. use of a pillow or knees flexed were recorded and repeated be the second examiner). The shoulder girdle is controlled in a neutral position before the shoulder is abducted to 110 degrees. Consequently, the wrist is extended and the forearm supinated before the shoulder is laterally rotated and the elbow extended. The range of elbow extension is measured via a goniometer strapped to the subject's medial arm with the fulcrum over the elbow joint. The test is terminated at the point of elbow extension where the patient reports verbally that they are experiencing "submaximal pain"(i.e. "a substantial discomfort, which corresponds with the greatest level of pain which the subject is prepared to tolerate, knowing that the test will be performed repeatedly"). The test is repeated with the neck side flexed away (i.e. contralateral) and towards (i.e. ipsilateral) the test arm.



A ULNT1: Starting position.

D ULNT1: Forearm supination.

B ULNT1: Shoulder abduction.

E ULNT1: Shoulder lateral rotation.

C ULNT1: Wrist extension.

F ULNT1: Elbow extension.

- **Definition of Abnormal Finding:**
  - Any one of the following criteria was classified as an "abnormal" or positive finding:
    - Reproduction of all or part of participant's symptoms.
    - Greater than 10 degree difference in elbow extension measured with goniometer
    - Contralateral neck side flexion increased or ipsilateral neck side flexion decreased client's symptoms.

- **Level of Measurement for Analysis:**
  - Nominal (positive or negative)

APPENDIX G- RECRUITMENT LETTER

Dear Sir/Madame:

We would like you to consider volunteering for a research study entitled: "*Reliability and diagnostic accuracy of clinical examination and patient self report measures in carpal tunnel syndrome*". The purpose of the study is to assess the usefulness of physical tests and questionnaires for the diagnosis of carpal tunnel syndrome.

You may qualify for the study if you:

- Have arm or hand symptoms (may or may not be carpal tunnel syndrome).
- Are 18 years or older.
- Are able to read and understand English.

What is involved?

- Completing paperwork (approximately 15-30 minutes)
- Undergoing physical tests by one or two Physical Therapists (30-60minutes)
- The paperwork and testing will take place on the same day immediately AFTER your scheduled time for electrodiagnostic testing at Saskatoon City Hospital.
- A parking reimbursement of $10 will be given to you for participating in the study.

If you are interested in participating in this study, please notify the Rehabilitation Department receptionist when you attend your appointment.

If you have any further questions, please contact:
Brenna Bath at 306 343 9984
or email: **bathb@shaw.ca**

APPENDIX H- PHYSICIAN LETTER

Date:

Re.:

Dear Dr. ,

Your patient consented to participate in a research study entitled: "*Reliability and Diagnostic Accuracy of Clinical Examination and Self-Report Measures in Carpal Tunnel Syndrome"*. As part of this study your patient completed questionnaires and underwent additional clinical examination procedures completed by a Physical Therapist.

Your patient agreed to have the results of this additional testing sent to you for your information.

The results were as follows:

Self Report Questionnaires:

- Disability of Arm Shoulder and Hand Questionnaire:
  Lower scores reflect less perceived disability and higher scores reflect more perceived disability.

- Carpal Tunnel Questionnaire:
  A lower score represents low reported symptom severity and higher perceived functional status.

Provocation Tests:

- Phalen's:
- Tinel's
- Carpal Compression Test:

Sensation Tests:

- Sharp/dull:
- Pressure sensation:

Strength Tests:

- Grip:

- Pinch
  - Tip to tip:
  - Key:
  - 3 jaw chuck:
- Thumb abduction:

Other tests:

- Flick sign:
- Thenar atrophy:
- Square wrist sign:
- Upper Limb Tension Test 1 (test for mobility and irritability of the median nerve):

If you have any questions or concerns, do not hesitate to contact me at
Ph: 343-9984 or email: **bathb@shaw.ca**.

Sincerely,

Brenna Bath BSc. PT, FCAMT, MSc (candidate)

APPENDIX I – CONSENT FORM

You are invited to participate in a study entitled: **Reliability and Diagnostic Accuracy of Clinical Examination and Self-Report Measures in Carpal Tunnel Syndrome**. *Please read this form carefully, and feel free to ask questions you might have.*

- **Name of researchers:**

Leonard Tan, Ph. D.	Brenna Bath, BSc PT
(Professor, Supervisor)	(Physical Therapist)
Dept. of Community Health and Epidemiology	MSc. Candidate CH&EP
University of Saskatchewan	University of Saskatchewan
Phone:	Phone: (306) 343-9984
	Email: bathb@shaw.ca

- **The purpose of the study:** To assess the reliability and validity of individual clinical examination items and self-report items for the diagnosis of carpal tunnel syndrome.

- **Potential Risks:** Some mild physical discomfort may be felt by the participants with some the physical tests; however, this risk will be minimized by physical screening done prior to the testing by a registered Physical Therapist. If a participant is found to have any pre-existing physical problems that would limit their safe participation in the study, they would not be allowed in the study.

- **Possible benefits of the study to the participant and others:** Nerve conduction tests (i.e. electrodiagnostic studies) are thought to be the best way to diagnose carpal tunnel syndrome. They can, however be time consuming, expensive to perform and uncomfortable patients. Clinical tests and questionnaires that accurately predictth epresence of carpal tunnel syndrome have the advantage of being relatively quick and easy to perform and could assist with earlier confirmation of a diagnosis. Earlier diagnosis could lead to earlier and likely better treatment in some cases. The results of the additional clinical and self-report tests performed for this study can be shared (with your consent) with your referring physician so that they may have more information at their disposal to help you with your condition.

- **The study procedures:** Study participants will be recruited from patients referred to Saskatoon City Hospital for electrodiagnostic testing due to arm, wrist or hand problems. As a participant, you will be asked to complete paperwork including information regarding your general health, age, education, income, work and hobbies. You will also be asked to complete questionnaires about your symptoms and functional abilities. The paperwork will take approximately 15-30 minutes to complete. You will then undergo a series of clinical tests performed by a registered Physical Therapist. The tests include carpal tunnel specific tests, strength and sensation measures and a test for the length of the median nerve (nerve in the arm). You may undergo the same series of clinical tests performed by a second registered Physical Therapist (this will be done with a random group from the study). The clinical tests will take approximately 15-30 minutes to complete. The results of your

electrodiagnostic testing will be used to compare to the results of the papaerwork and additional clinical testing. You will receive $10 to reimburse you for your time and parking. This parking honorarium will also be paid to participants who choose to withdraw upon attending their study appointment.

- **Confidentiality:** Confidentiality of study participants will be kept by assigning subject numbers to each subject. Only the subject's number will be present on paperwork. The key linking participant name and subject number will be kept in a secure location onsite at the Physical Medicine and Rehabilitation Department at Saskatoon City Hospital and will be destroyed upon completion of the data collection period. Although the data from this study may be published and presented at conferences, the data will be reported in group form, so that it will not be possible to identify individuals. The consent forms will be stored separately from the results of the testing so that it will not be possible to associate a name with any given set of responses.

- **Right to Withdraw:** You are free to withdraw from the study at any time. Withdraw from the study will not affect your academic or employment status, and/or access to, or continuation of, services provided by public agencies such as the University, hospitals, or social services. Your access to treatment will not be affected by your decision to participate. If you decide to withdraw, you may choose to have their data deleted from the study and destroyed.

- **Researcher decisions and interpretations:** If for any reason the researcher decides that involvement by a participant is not needed or not appropriate, the researcher can choose to discontinue the participant's involvement in the study, in which case the participant's data will be deleted from the study and destroyed.

- **Reports and presentations:** The information obtained from this study will be used in a number of different forms. The information will be used in a master's thesis and may be used in articles published in medical journals and in oral or poster presentations to groups of researchers and students.

- **Questions:** If you have any questions concerning the study, please feel free to ask at any point; you are also free to contact the researchers at the numbers provided above if you have questions at a later time. This study has been approved on ethical grounds by the University of Saskatchewan Behavioural Sciences Research Ethics Board on August 19th, 2004. Any questions regarding your rights as a participant may be addressed to that committee through the Office of Research Services (966-2084). Out of town participants may call collect.

Consent to Participate:

I, _____ (name), hereby acknowledge that:

- I have read and understood the description provided above.
- I have been provided with an opportunity to ask questions and my questions have been answered satisfactorily.
- I consent to participate in the study described above, understanding that I may withdraw this consent at any time.
- I agree that the researchers may access the results of my electrodisgnostic testing.
- I have received a copy of the consent form for my own records and,
- I would / would not (please circle one) like a summary of the additional test results sent to my referring physician.

_____ _____

Participant Date

_____ _____

Researcher Date

APPENDIX J- PARTICIPANT GENERAL PAPERWORK

Thank you for your agreement to participate in this study. Please answer the following questions to the best of your knowledge. To ensure confidentiality, please do not put your name on any of the following pages. If you have any questions about the questionnaires, please ask to speak to the study co-ordinator.

About You:

1. Age: _____

2. Gender:
   Male
   Female

3. Current Marital Status (please circle):
   Married
   Separated
   Divorced
   Widowed
   Never Married

4. Education:
   Did not complete Grade 12
   High School
   Trade School
   Some University
   University Degree
   Graduate Degree

5. Family Income:
   <$15,000
   $16,000-29,000
   $30,000-59,000
   $60,000-99,000
   Equal to or greater than $100,000

6. Height: _____

7. Weight: _____

8. Whish side is your dominant hand?
   Right Left Both

About Your Job:

9. Please check your main form of work
   - paid work-full time
   - paid work-part time
   - unemployed
   - housework
   - disabled
   - student
   - retired

10. If employed, what is your occupation: _____

11. If employed, does your job involve any of the following?
   - repetitive hand or wrist movements
   - work in cold environments
   - vibration through hands and wrists

12. Are you currently receiving benefits through the Worker's Compensation Board (WCB) for your hand/wrist/arm problem? Yes No

13 Are you currently receiving benefits through Saskatchewan Government Insurance (SGI) for your hand/wrist/arm problem? Yes No

About your medical status:

14. Please check any of the following medical conditions that you may have or have had:
   - Diabetes
   - Rheumatoid Arthritis
   - Osteoarthritis
   - Fractured wrist
   - Thyroid condition
   - Other: _____

15. Do you smoke? Yes No

About your symptoms:

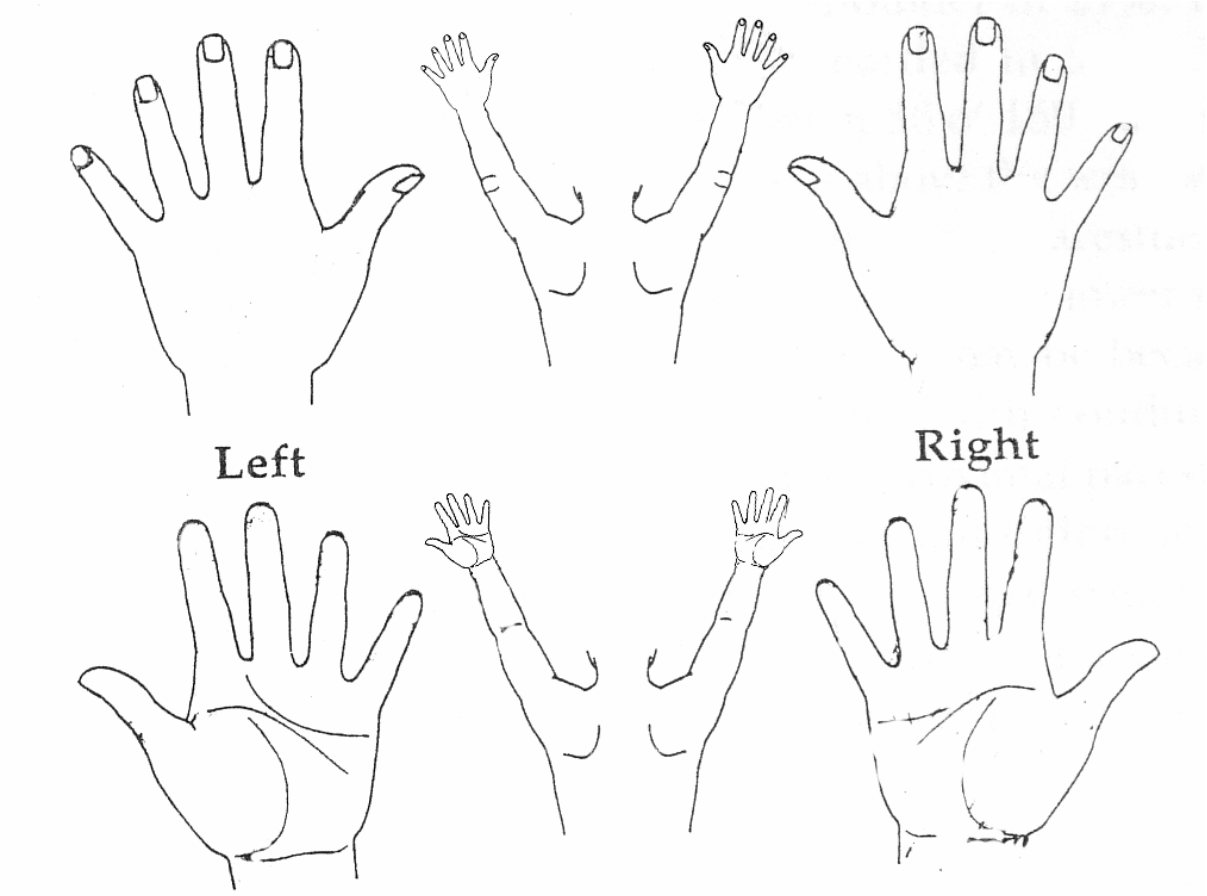
16. Which side do you have symptoms with your hand/wrist/arm?
   Right Left Both

* If both sides bother you, which side is worse?

Right Left

17. How long have you had symptoms at your hand/wrist/arm? _____

18. HAND/ARM DIAGRAM: Please shade in where you feel pain, numbness, tingling or burning sensations in your hands/arms.



APPENDIX K
Physical Exam Reporting Form: CTS Study

Date: _____ **Subject Number: _____**
Examiner Initials: _____

Provocation Tests:

1. **Phalen's:** Right: ____ positive ____negative
   Left: ____ positive ____negative

2. **Tinnel's:** Right: ____ positive ____negative
   Left: ____ positive ____negative

3. **Carpal Compression:** Right: ____ positive ____negative
   Left: ____ positive ____negative

Sensation Tests:

1. **Sharp/Dull:** Right: ____ positive ____negative
   Left: ____ positive ____negative

2. **Pressure:** Right: ____ positive ____negative
   Left: ____ positive ____negative

Strength Tests:

1. **Grip** (kg): Right: trial 1:_____ trial 2:_____ trial 3:_____
   Left: trial 1:_____ trial 2:_____ trial 3:_____

2. **Pinch** (kg):
   a. **Tip to tip**: Right: trial 1:_____ trial 2:_____ trial 3:_____
      Left: trial 1:_____ trial 2:_____ trial 3:_____

   b. **Key**: Right: trial 1:_____ trial 2:_____ trial 3:_____
      Left: trial 1:_____ trial 2:_____ trial 3:_____

   c. **3 jaw chuck**: Right: trial 1:_____ trial 2:_____ trial 3:_____
      Left: trial 1:_____ trial 2:_____ trial 3:_____

3. **Thumb abduction**: Right: Grade (1-5) _____ Left: Grade (1-5) _____

Other Tests:

1. **Flick sign:** Right: ____positive ____negative Left: ____positive ____negative

2. Thenar atrophy: Right: ____positive ____negative Left: ____positive ____negative

3. **Square wrist** (mm): Right: A/P:_____ M/L :_____ Left: A/P:_____M/L :_____

4. **ULNTT 1**:

Right: elbow ext ROM_____
____ pt's symptoms reproduced
____>10deg difference elbow extension
____contralateral neck SF increased
or ipsilateral neck SF decreased symptoms

Left: elbow ext ROM_____
____ pt's symptoms reproduced
____>10deg difference elbow extension
____contralateral neck SF increased
or ipsilateral neck SF decreased symptoms

APPENDIX L – CTS DIAGNOSTIC CRITERIA AND EDX PROCEDURES

CTS Diagnostic Criteria for EDX:

- Positive EDX for CTS includes any one of the following:
  - Transpalmar difference between ulnar and median nerves of greater than 0.3 ms peak latency on mixed nerve study
  - Antidromic sensory conduction velocity of less than 50m/s +/- amplitude less than 10 uV or absent sensory nerve action potential
  - Motor distal latency of >4.2 ms +/- amplitude of less than 5 mV.

Electrodiagnostic Procedures:

- All EDX testing will be performed by a physiatrist.
- Bilateral median and ulnar motor and sensory nerve conduction testing will primarily be used in the assessment of suspected CTS.
- Electromyography (EMG) will typically not be used for the assessment of CTS and ulnar neuropathy unless severe neuropathy is suspected.
- EMG will likely be used for the assessment of radiculopathy, brachial plexopathy and other upper extremity neuropathies (i.e. radial neuropathy).
- EMG assessment involves needle insertion into key muscles and evaluation of insertional and spontaneous activity as well as motor unit recruitment.

APPENDIX M
CTS STUDY: EDX RESULTS

EDX results

- *Positive EDX for CTS includes any one of the following:*
  - *Transpalmar difference between ulnar and median nerves of greater than 0.3 ms peak latency*
  - *Antidromic sensory conduction velocity of less than 50m/s +/- amplitude less than 10 uV*
  - *Motor distal latency of >4.2 ms +/- amplitude of less than 5*

Right:	Positive	Negative	
	Mild	Moderate	Severe
Left:	Positive	Negative	
	Mild	Moderate	Severe

Differential Diagnosis (circle any / all that apply)

CTS

Ulnar Neuropathy

Radial Neuropathy

Brachial Plexopathy

Cervical Radiculopathy

Diabetic Neuropathy

Other: _____

Unknown

APPENDIX N
SPSS Logistic Regression Results Tables

General Classification Group:

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1(a)	sss6cat(1)	22.812	17974.852	.000	1	.999	80773773 39.791
	Constant	-21.203	17974.852	.000	1	.999	.000
Step 2(b)	ulttsymp1(1)	1.856	1.102	2.836	1	.092	6.400
	sss6cat(1)	23.968	17571.092	.000	1	.999	25664273 747.971
	Constant	-22.829	17571.092	.000	1	.999	.000

a Variable(s) entered on step 1: sss6cat.
b Variable(s) entered on step 2: ulttsymp1.

Restricted Classification Group:

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1(a)	sss6cat(1)	2.485	1.118	4.940	1	.026	12.000
	Constant	-1.792	1.080	2.752	1	.097	.167
Step 2(b)	sss6cat(1)	2.091	1.152	3.295	1	.069	8.095
	ctqsslev(1)	1.170	.619	3.569	1	.059	3.222
	Constant	-2.230	1.141	3.820	1	.051	.108

a Variable(s) entered on step 1: sss6cat.
b Variable(s) entered on step 2: ctqsslev.